

**EXPERIMENTAL TIME-DOMAIN CONTROLLED SOURCE
ELECTROMAGNETIC INDUCTION FOR HIGHLY CONDUCTIVE
TARGETS DETECTION AND DISCRIMINATION**

A Dissertation

by

ALFONSO BENAVIDES IGLESIAS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

May 2007

Major Subject: Geophysics

**EXPERIMENTAL TIME-DOMAIN CONTROLLED SOURCE
ELECTROMAGNETIC INDUCTION FOR HIGHLY CONDUCTIVE
TARGETS DETECTION AND DISCRIMINATION**

A Dissertation

by

ALFONSO BENAVIDES IGLESIAS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Mark E. Everett
Committee Members,	Richard Carlson
	William W. Sager
	Steven Wright
Head of Department,	John H. Spang

May 2007

Major Subject: Geophysics

ABSTRACT

Experimental Time-domain Controlled Source Electromagnetic Induction for Highly
Conductive Targets Detection and Discrimination. (May 2007)

Alfonso Benavides Iglesias, B.S; M.A.S., Universidad Central de Venezuela

Chair of Advisory Committee: Dr. Mark E. Everett

The response of geological materials at the scale of meters and the response of buried targets of different shapes and sizes using controlled-source electromagnetic induction (CSEM) is investigated. This dissertation focuses on three topics; i) fractal properties on electric conductivity data from near-surface geology and processing techniques for enhancing man-made target responses, ii) non-linear inversion of spatiotemporal data using continuation method, and iii) classification of CSEM transient and spatiotemporal data.

In the first topic, apparent conductivity profiles and maps were studied to determine self-affine properties of the geological noise and the effects of man-made conductive metal targets. 2-D Fourier transform and omnidirectional variograms showed that variations in apparent conductivity exhibit self-affinity, corresponding to fractional Brownian motion. Self-affinity no longer holds when targets are buried in the near-surface, making feasible the use of spectral methods to determine their presence. The difference between the geology and target responses can be exploited using wavelet decomposition. A series of experiments showed that wavelet filtering is able to separate target responses from the geological background.

In the second topic, a continuation-based inversion method approach is adopted, based on path-tracking in model space, to solve the non-linear least squares problem for unexploded ordnance (UXO) data. The model corresponds to a stretched-exponential decay of eddy currents induced in a magnetic spheroid. The fast inversion

of actual field multi-receiver CSEM responses of inert, buried ordnance is also shown. Software based on the continuation method could be installed within a multi-receiver CSEM sensor and used for near-real-time UXO decision.

In the third topic, unsupervised self-organizing maps (SOM) were adapted for data clustering and classification. The use of self-organizing maps (SOM) for central-loop CSEM transients shows potential capability to perform classification, discriminating background and non-dangerous items (clutter) data from, for instance, unexploded ordnance. Implementation of a merge SOM algorithm showed that clustering and classification of spatiotemporal CSEM data is possible. The ability to extract target signals from a background-contaminated pattern is desired to avoid dealing with forward models containing subsurface response or to implement processing algorithm to remove, to some degree, the effects of background response and the target-host interactions.

To Delia and Irene, in the present and for the future

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to my advisor, Mark Everett, for his guidance and encouragement during my graduate studies and in the preparation of this dissertation. He also shared with me an invaluable friendship and advice which will be useful in my future professional life.

I also thank the members of my advisory committee, William Sager (who was my teacher during two semesters), Richard Carlson, Steven Wright for their support and helpful comments, and Fred Chester for taking a valuable time in reviewing this manuscript before the final examination. Lloyd Morris for his support in site selection and data acquisition at Riverside Campus and George Robitaille (Aberdeen Proving Ground) for providing us with a standardized UXO kit.

My sincere gratitude to CDCH-UCV (Consejo de Desarrollo Científico y Humanístico, Universidad Central de Venezuela) for sponsoring part of my PhD studies. Also, Texas A&M University Sponsored Student Programs crew: Violetta Cook, Angela Sánchez, Nancy Barnes, and Margit Garay at International Student Services for helping me with all the necessary paperwork to “preserve” my immigration status.

I am also grateful to Carl Pierce Jr. for helping me during the field work and specially for his friendship. Thanks to him, I have learned a lot about the American culture and Native American heritage. Also, I wish to thank friends and office mates during my stay: Javier Pérez (PDVSA-Intevep), Douglas Sassen, Joshua King, Zach Long, Jamie Collins, David Dickins, Souvik Mukherjee, Jakub Velimsky (now at the Charles University in Prague), Leticia Espinosa, and my colleagues and friends in Caracas: Rommel Whilchy, Armando Sena, Liliana López and Cruz Ramón Blanco for their support and helping me keeping things running smoothly despite the distance.

My gratitude to Louise Pellerin, George Jiracek, Erika Gasperikova, Les Beard

and the anonymous reviewers for helping me during the reviewing process of the papers published during my graduate program which constitutes the core of the present dissertation, Barbara Hammer from the Institute of Computer Science, Clausthal University of Technology in Germany for her support and suggestions about merge self-organizing maps.

Also, I want to thank the Geology and Geophysics Department staff: Michele Beal, Debbie Schorm, Debra Stark and Gwen Tennell; they certainly helped me in crucial moments I needed the most.

Finally, to my wife Delia Perez-Nunez for her support and patience during this long journey.

This work was partially supported by U.S. DOD grant SERDP UX-1312.

TABLE OF CONTENTS

CHAPTER		Page
I	INTRODUCTION	1
	1. Experimental observations	3
	A. Geological noise	3
	B. CSEM data acquisition and highly conductive targets	4
	C. The general CSEM problem for conductive targets	8
	2. Modeling EM induction in conductors	9
	A. Models for conductive targets of high symmetry	9
	B. Physics-based empirical models	10
	3. Dissertation structure	11
II	EXPERIMENTAL CHARACTERIZATION OF CSEM RE- SPONSES FROM GEOLOGY AND CONDUCTIVE TARGETS	13
	1. Introduction	13
	2. Methods	17
	3. Results and discussion	23
	A. Geological noise	23
	B. Target-background separation	32
III	INVERSION OF SPATIOTEMPORAL CSEM DATA US- ING A CONTINUATION METHOD	39
	1. Introduction	39
	A. The inverse problem	40
	B. Continuation methods	40
	C. Outline	41
	2. Methods	42
	A. Inversion	42
	B. Forward response	45
	C. Data acquisition	47
	3. Results and discussion	49
	A. Synthetic data	50
	a. Sensitivity to initial model	52

CHAPTER		Page
	b. Convergence and comparison with LM algorithm	54
	B. 57mm M22 antitank ordnance	54
	C. 40mm MKII projectile	56
	D. 60mm M49A3 mortar	56
IV	UNSUPERVISED CLASSIFICATION OF CSEM DATA USING SELF-ORGANIZING MAPS	61
	1. Introduction	61
	A. Feed forward neural networks	63
	B. Self organizing maps	65
	C. Classification of time-domain CSEM responses	67
	D. Chapter organization	68
	2. Methods	68
	A. Kohonen SOM	69
	a. SOM visualizations	73
	B. SOM for time series	75
	a. Generalized recurrent SOM	76
	C. Merge SOM	79
	a. MSOM algorithm	80
	b. Interpretation of MSOM learning	83
	D. Datasets and pre-processing	84
	a. Benchmark datasets for SOM algorithm	84
	b. Benchmark datasets for MSOM algorithm	88
	c. Geonics EM63 data	89
	3. Results and discussion	95
	A. Benchmarking of SOM and MSOM algorithms	95
	B. SOM for EM63 central-loop transients	102
	C. MSOM for EM63 spatiotemporal patterns	104
V	CONCLUSIONS	112
	REFERENCES	115
	APPENDIX A	130
	VITA	133

LIST OF FIGURES

FIGURE	Page
1 Effect of a moving TX-RX system over a buried horizontal aluminum plate	6
2 Spatiotemporal response of a 81 mm M170 mortar	7
3 Texas A&M Riverside Campus with the site positions used in the present work.	18
4 General map of lake site in south Texas	20
5 EM31 Experimental setup	22
6 EM31 apparent conductivity profiles from the pipe tests shown in Figure 5	24
7 Results from EM31 and EM34 apparent conductivity profiles for site 1	26
8 Wavelet approximation and details for the average response of a type-1 pipe in parallel array configuration using DWT	28
9 Wavelet filtering of the EM31 profiles	29
10 Target-free EM31 apparent conductivity maps for the site 3	30
11 2-D omni directional log variograms from site 3 EM31 data (see Figure 10)	31
12 2-D FT for the P-array (left) and T-array (right) orientation conductivity maps	32
13 EM31 apparent conductivity maps in the presence of pipes	33
14 Log-variograms for the P (left) and T (right) array apparent conductivities of Figure 13	34

FIGURE		Page
15	Composite $C(x) = T(x) - P(x)$ for the pipe tests in site 1	35
16	Controlled EM31 buried pipe test	36
17	EM34 profile data and processing results at the lake site using DWT	37
18	Schematic of a continuation path Γ in a bounded model parameter search space.	45
19	Schematic of the experimental multi-receiver data acquisition setup using a modified Geonics EM63	48
20	Spatiotemporal data representation	49
21	Inversion results for a synthetic dataset corresponding to a stretched- exponential decaying magnetic spheroid	51
22	Continuation path projections for different starting models using the synthetic dataset in Figure 21	53
23	Inversion results for a 57mm M22 antitank projectile	57
24	Inversion results for a 40mm MKII projectile	58
25	Inversion results for a 60mm M49A3 mortar	60
26	Visualization of convex regions in a SOM	69
27	Sketch of a self-organizing map learning process	71
28	Sketch of a merge self-organizing map learning process	82
29	Two concentric Gaussian clusters dataset	85
30	Datasets from synthetic transient data	86
31	Dataset of spatial dipole responses	87
32	Some data projections of the Iris dataset	87
33	Two-cluster synthetic spatiotemporal data	89
34	Three-cluster synthetic spatiotemporal data	90

FIGURE	Page
35	EM63 transients dataset 91
36	Target burial orientations 93
37	Geonics EM63 spatiotemporal responses 94
38	SOM results for two concentric Gaussian distributions data 96
39	SOM results for 2-cluster of transient data 97
40	SOM results for 3-cluster of transient data 97
41	SOM results for 4-cluster of dipole responses 98
42	U* matrix and map activation for the Iris dataset 99
43	MSOM results for 2-cluster synthetic spatiotemporal data 100
44	MSOM results for 3-cluster synthetic spatiotemporal data 101
45	SOM results for EM63 central-loop transient data 103
46	Average cost function per epoch and U*-matrix for EM63 spa- tiotemporal data 104
47	MSOM activation paths for the EM63 spatiotemporal background data 105
48	MSOM activation paths for EM63 spatiotemporal UXO data in orientation C3 106
49	MSOM activation paths for EM63 spatiotemporal clutter data in orientation C1 107
50	MSOM activation paths for EM63 spatiotemporal clutter data in orientation C4 107
51	Activation of the MSOM to a background input pattern. 109
52	MSOM activation paths for EM63 spatiotemporal UXO data in orientation C5 109
53	Activation of the MSOM to a clutter input pattern. 111

CHAPTER I

INTRODUCTION

Unexploded Ordnance (UXO) detection, discrimination and removal represent a priority environmental and geotechnical problem in many regions of the world including Central America, Africa, and Asia. In the U.S., UXO cleanup is driven by a congressional mandate for safe base re-alignment and closure (BRAC) as military ranges are made sustainable or transferred to civilian use. The common factor in cleanup operations is the huge final cost per UXO removed, mainly associated with unnecessary digging operations. The probability of UXO detection on documented test sites can exceed 90% in carefully executed geophysical surveys; however, the false alarm rates (the number of non-UXO targets that must be excavated for each UXO item found) remain quite high (from tens to hundreds depending on site specific conditions). To appreciate the impact of high false alarm rates, consider a \$30 million cleaning operation of a 5,000 acre heavily contaminated site (Butler, 2003); only 9% of the cost corresponds to UXO removal while 76% is for non-UXO removal (the remaining 15% includes site conditioning and other expenses). The existence of tens of thousands of contaminated sites (only in the U.S.) turns the UXO cleaning problem into a multi-billion dollar activity. Of the three problem components (detection, discrimination and removal), discrimination is the critical one; it enables the geophysicist to identify which detected targets are certainly UXO, and it generates complementary information about UXO type, depth and orientation; which is valuable for removal operations.

In UXO discrimination, the geophysicist faces a small scale problem in which

This dissertation follows the style of the journal *Geophysics*.

the host materials may be heterogeneous (a mature soil profile, a highly weathered rock sequence, sediments, plus potentially many man-made disturbances in the form of metallic clutter, dig-and-backfilled volumes, etc.). There is one or more highly conductive targets embedded in the subsurface at depths ranging from centimeters to several meters. Controlled-source electromagnetic induction (CSEM¹) techniques are useful to solve UXO problems because: i) there are a large conductivity contrast between target and host ($10^6 - 10^7$ S/m for targets and clutter compared with $10^{-3} - 10^{-1}$ S/m for near-surface backgrounds); and ii) the response does not require the target to be ferromagnetic. The target is a source of induced currents whose spatiotemporal response contains information about its electromagnetic and geometric properties.

The objective of this dissertation is to develop UXO classification tools using time-domain controlled-source electromagnetic induction. Experimental design, data acquisition, physics-based non-linear model inversion, the creation of a target response feature library and a feature-based classifier, are the essential tasks. The key contributions of this dissertation can be summarized as follows: i) definition of a new mode of CSEM data acquisition through the use of multi-receiver data; ii) construction of a library of actual and synthetic target responses using several sensor-target configurations; iii) implementation of non-linear inversion algorithms using simple physics-based models and numerical continuation methods; iv) design and testing of a feature-based classifier.

¹Remote sensing technique which uses electric or magnetic sources (transmitter, TX) to induce currents in the subsurface. Single or multi-frequency harmonic TX waveforms are used in frequency-domain systems while square or triangular waveforms are used on time-domain systems.

1. Experimental observations

In comparison with traditional geophysical resource prospecting, near-surface geophysics does not require large arrays of sensors, high power sources, and long periods of time for data to be acquired. However, assumptions such as homogeneity of the physical properties, far-field approximations, and negligible effects of the topography may become the main sources of error, and they are dependent on the survey site, the instrument used, and the way the survey is conducted.

Within the near surface, there are several factors that affect geological material properties: weathering processes, degree of water saturation, differential stresses, etc. There are also physical properties such as texture, porosity, bedding, and lateral/vertical compositional variations. These factors and properties shape the CSEM response.

A. Geological noise

It has been shown that geological heterogeneities develop in patterns that are essentially the same at different length scales. This self-similarity can be described mathematically by saying that the number of features of some length L scales as a negative power of L . This is a basic description of a fractal (Mandelbrot, 1983). Topography and ore tonnage, for example, exhibit a fractal dimension (Turcotte, 1992). The same result is obtained for porosity in some sedimentary rocks (Katz and Thompson, 1985).

Recently, it has been shown that CSEM spatial profiles of apparent conductivity, using single frequency instruments, exhibit a power-law Fourier wavenumber spectrum whose fractional exponent corresponds to a fractal (Everett and Weiss, 2002). In chapter II we show that like topography, conductivity maps over areas of a few hun-

dred square meters show similar characteristics contrasting with the smooth, localized response of near-surface conductive targets such as pipes and cased boreholes. This dichotomy between fractal geological noise and deterministic target signals makes possible the application of wavelet-based filtering techniques to recover target response from background noise.

B. CSEM data acquisition and highly conductive targets

Acquisition of time-domain CSEM responses is typically carried out using a pair of co-axial transmitter (TX) and receiver (RX) coils. The coils are suspended a few decimeters above the ground surface by using wheels or hand-held devices and moved in tandem. The TX waveform is an alternating square wave with a rapid (tens to hundreds of μs) linear ramp-off. Shortly after the ramp-off, the RX coil is activated, sampling the time variation of the secondary magnetic field from induced currents in the target using logarithmic or linear spaced time gates. There is a large literature regarding the use of these systems for UXO detection and characterization (McNeill, 1980; McNeill and Bosnar, 2000; Geonics, 2001; Snyder et al., 2003). The simple TX-RX geometry employed in commercial systems is convenient for fast reconnaissance surveys but it is not optimal for recording high quality target responses. The reason is that, for each measurement, both TX and RX are moved with respect to the target, producing a map that contains target response under variable excitation conditions resulting in a target response that is more difficult to interpret. Figure 1 shows 26 profiles corresponding to the EM63 time channels for a horizontally buried Aluminum plate. Notice that the profiles are completely different in shape and amplitude when the TX loop is moved above the target. In a fixed central-loop Rx configuration the recorded transient would have been the data corresponding to station number 12. In a multi-receiver array, the response would correspond to the top plot of Figure 1.

A multi-RX array with fixed TX location generates an informative spatiotemporal map of the CSEM response which preserves information about the relative geometry between TX and target. CSEM responses from a large variety of isolated conductive targets exhibit two common characteristics. First, the shape of the spatial response remains basically the same over the entire transient decay; each spatial profile is a scaled version of the others. The scale factor from one time channel to the next one follows a non-linear trend. This sort of self-similarity is easier to observe at the late-time channels of the decay, when the signal from the background has largely vanished. Second, the nodes of the spatial response (RX locations where the measured time variation of the secondary field component is zero) remain at the same location during the whole decay. This feature is not observed at early-time channels when a pure background transient is recorded (see Figure 2). Based on the above observations, spatiotemporal CSEM responses for isolated, highly conductive targets $V(x, t)$ can be written as the product of spatial function $X(x)$ and a time-dependent function $T(t)$,

$$V(x, t) = X(x)T(t) \quad (1.1)$$

This separable equation holds at late-time responses (Everett et al., 2005) and it can be extended to intermediate and early times if the contribution from the background signal is subtracted properly. Equation 1 has been tested using single and composite targets whose size is smaller than the TX coil.

the spatial function $X(x)$ can generally be described using a point magnetic dipole while the time function $T(t)$ can be viewed as a non-exponential decay signal (the superposition of several fundamental decay processes). The most general property that apparently holds for all kinds of targets is the single exponential decay mode for late-times, associated with the slowest eddy current mode that persists in the bulk of the target body (the term late-time used here is related to the fundamental target

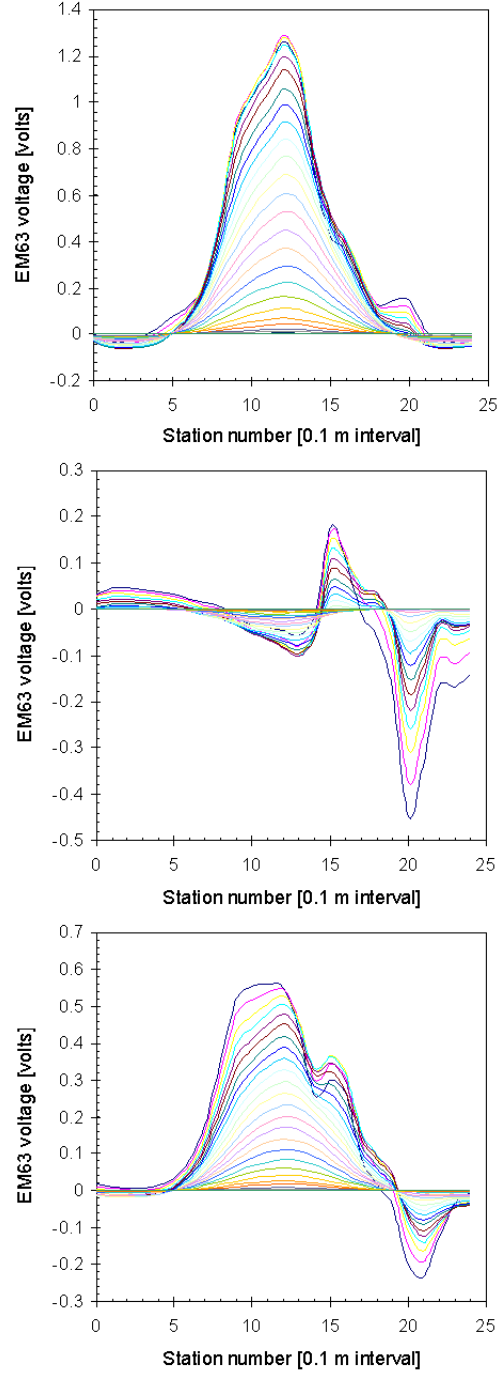


Fig. 1. Effect of a moving TX-RX system over a buried horizontal aluminum plate. The plate ($0.3 \times 0.3 \text{ m}$) is buried 0.4 m at station number 12. There are 26 profiles on each plot corresponding to the EM63 time channels. Top: TX loop (dimensions $0.3 \times 0.3 \text{ m}$) centered above the plate. Middle: TX loop has been moved 0.5 m to the right. Bottom: TX loop moved an additional 0.3 m to the right.

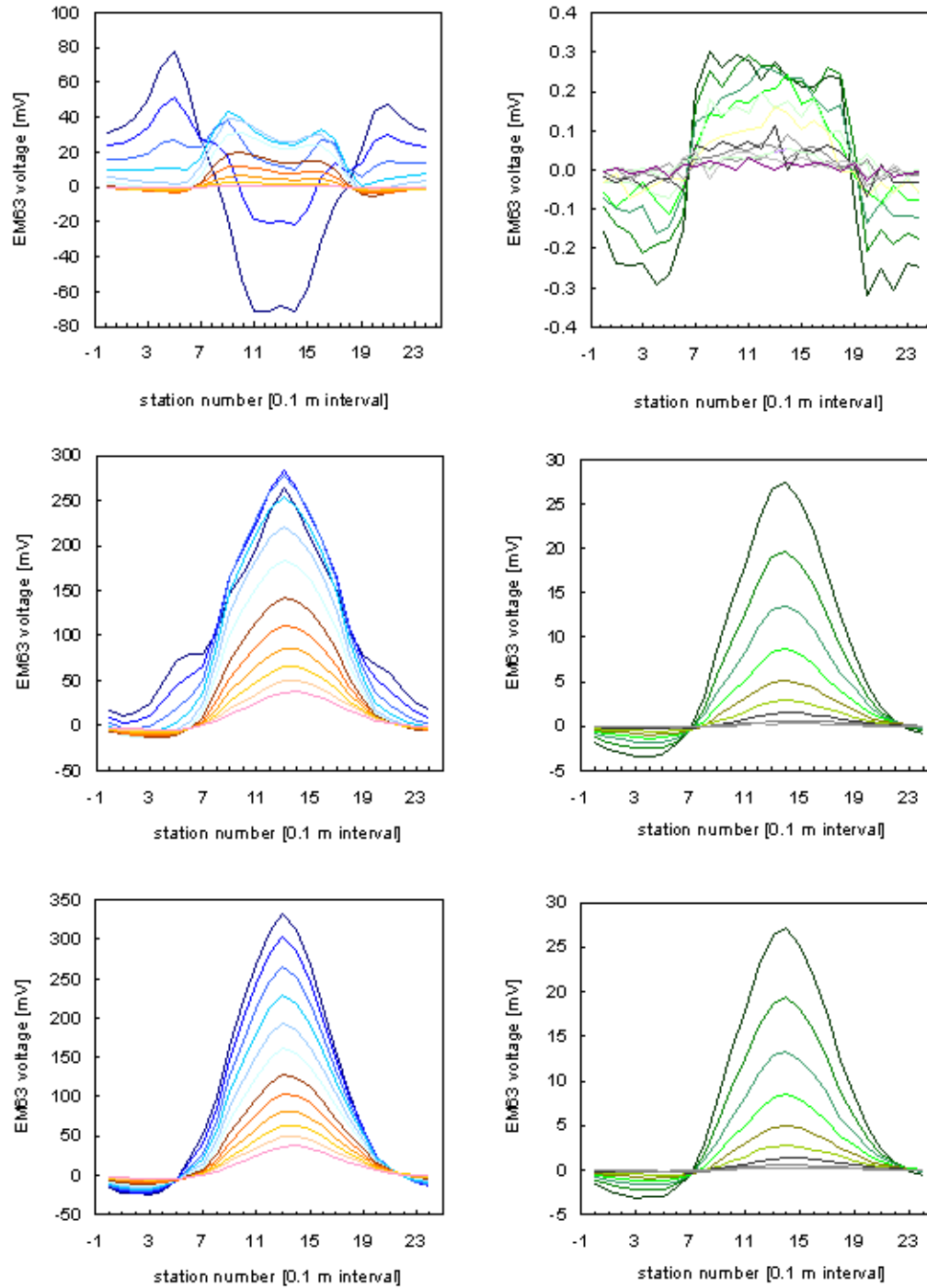


Fig. 2. Spatiotemporal response of a 81 mm M170 mortar. The ordnance is buried horizontal at $0.4m$ depth. Plots in the left and right columns comprise the first and last 12 EM63 time channels, respectively. Top row: Background response; middle row: background and target response; bottom row: background-subtracted response.

response and it is not necessarily linked to the last acquisition channels of commercial instruments). Some empirical functions have been used for describing the CSEM transient using just a few parameters (Pasion and Oldenburg, 2001; Everett et al., 2005). No reports about time decay rate distribution analysis for time-domain CSEM data have been found in the literature despite its success for classification purposes in many fields of applied physics and chemistry (Istratov and Vyvenko, 1999).

C. The general CSEM problem for conductive targets

UXO's are composite metal targets with a wide range of geometries. Generally, iron is the principal component but metals such as aluminum, copper, brass (copper-tin alloy), hardened steel, and titanium constitute complementary components including the nose, fins, rotation ring, and fuse. The manner in which induced currents are established and evolve in time and space depends on the electrical conductivity, magnetic permeability, shape and orientation of each UXO part, electric contacts between adjacent parts, and mutual electromagnetic coupling.

CSEM response measurements of different target parts and background responses separately (keeping the same locations as on the whole ordnance), provides the interesting result that superposition of tip, cylindrical body and background signals is greater than the response of the whole ordnance in the background. Thus, targets with several parts generate a single spatiotemporal response that is a complicated function of the individual parts plus interaction terms.

The existence of ferrous and non-ferrous materials further complicates the physical description of the target CSEM response. During and after the TX turn-off in a time-domain CSEM experiment, eddy currents in non-ferromagnetic conductors will circulate in such a direction, according to the Lenz's law, as to produce fields that oppose the vanishing incident field. Amperian currents induced in ferrous material,

however, provides stronger induced fields because the magnetic field inside the material is stronger than in the case of a non-ferromagnetic material (Saslow, 1991; Baum, 1999). For axial-symmetric targets with a dominant principal axis, a non-ferrous and ferrous response is expected when the TX magnetic field is normal or it is aligned to the target’s principal axis, respectively (Saslow, 1991; McNeill and Bosnar, 2000). Realistic CSEM modeling should take into account geometry, electric/magnetic properties, location and orientation of the target with respect to the sensor.

2. Modeling EM induction in conductors

Geophysical literature regarding modeling the inductive response of conductive targets with specific geometries is limited. There are several reasons; geologic targets of interest for energy and mining exploration are composed of stacks of layers, faults, dikes, and most of them are not more conductive than the surrounding rocks. One exception is modeling ore deposits such as massive sulphide complexes due to its higher electrical conductivity with respect to the host rock. In the case of geotechnical or environmental studies the situation is different; detection and discrimination of metallic objects is an important task. From rebar, pipe location and integrity assessment to unexploded ordnance discrimination, there is a wide range of applications that requires some degree of modeling. In this section a brief review about analytic solutions for simple target shapes and more general, but limited, physics-based empirical models will be reviewed.

A. Models for conductive targets of high symmetry

The literature on analytic solutions for the CSEM response is mostly restricted to solid spherical or spheroidal shapes calculated in the quasi-static regime; displacement

currents in the target and conduction-displacement currents in the host are neglected.

Several analytic solutions can be found in the literature for simple, symmetric targets in both frequency and time domain. In the frequency-domain, for a conductive sphere shielded by a concentric conductive shell of larger diameter in the whole space (Fuller, 1971); a hollow sphere excited by an arbitrary dipole (Mrozynsky, 1998), and a solid, permeable spheroid with arbitrary, harmonic excitation (Ao et al., 2002). The hollow sphere is a suitable starting model for symmetric UXO because it constrains the radial eddy current distribution inside the conductor and gives more realistic behavior at intermediate and late-times. Time-domain solutions can also be found for two concentric layers with different conductivities, permeability and dielectric constants embedded in a conductive whole space (Nabighian, 1971), and for a permeable and conductive sphere in a conductive infinite space with an arbitrarily located dipole source (Singh, 1973).

Regarding secondary magnetic field asymptotic regimes, the early time stage is characterized by an essentially infinite number of decays which sometimes superimpose to generate a $t^{-1/2}$ power-law decay in the external magnetic field (Weichman, 2003). On the other hand, a generic study in frequency and time-domains for the magnetic response of conducting bodies in an empty whole space (Kaufman, 1978) states that the late-time secondary magnetic field decay is single exponential. The transient field is more sensitive to changes in conductivity than the frequency domain methods, and transient methods have better classifying properties than amplitude-phase methods.

B. Physics-based empirical models

Pasion and Oldenburg (2001) describe CSEM transient responses as the sum of two orthogonal independently decaying dipoles. The temporal function $T_i(t)$ for each

dipole is an empirical non-exponential decay of the form

$$T_i(t) = \kappa_i(\alpha_i + t)^{-\beta_i} \exp(t/\tau_i). \quad (1.2)$$

The ratios of the inverted parameters κ_1/κ_2 and β_1/β_2 are target geometry indicators (plate-like or rod-like) for ferrous targets. A similar approach is presented by Das et al. (1990) in which the induced magnetic dipole for a prolate spheroid is expressed as a function of the primary magnetic field, direction of the principal direction of the target and the longitudinal and transversal magnetic polarizabilities. The use of several decaying dipole sources has been studied for composite targets in the frequency-domain (Zhang et al., 2003).

Empirical models consisting of two orthogonal and independently decaying magnetic dipoles approximate very well the response of targets whose size is comparable with the region where the primary magnetic field is reasonably homogeneous. Currently, the size of the TX coil in commercial units ranges from 0.2 to 1 m, while ordnance lengths vary from a few centimeters to a few meters; in this scenario it is difficult to assess the validity of point sources as phenomenological models to represent targets with sizes comparable to the TX source or whether a more general model is required.

3. Dissertation structure

This presentation follows the order the research was carried on. Each chapter contains partial results published, in press, or under review about the topics mentioned on this chapter: background geological noise, physics-based model data interpretation, and classification. A better understanding of near-surface electromagnetic response characteristics can lead to improvements in incorporating those features into models,

or to remove them properly, giving rise to simpler EM responses for inversion. In chapter II, a complementary study of apparent conductivity profiles and an extension to 2-D conductivity maps is presented, providing some experimental determinations of self-affinity, spectral content, and the effects of man-made targets.

Chapter III deals with non-linear inversion using the response of magnetic spheroids as the forward model. The application of a continuation method to control the inversion process leads to a wider range of initial models and a more robust inversion. Examples provided came from the experiments carried out using a modified Geonics EM63, a time-domain metal detector. Technical data can be found at appendix A.

Chapter IV comprises the latest research to come up with a flexible method to interpret Geonics EM63 data aiming for discrimination of background and clutter signals from those produced, for instance, by unexploded ordnance. The paradigm of artificial neural networks was chosen using an unsupervised learning strategy. The chapter describes the use of self-organizing maps for static input patterns and a recent recurrent self-organizing map for time series.

In this way, this dissertation covers three basic aspects of near-surface geophysics and provides some insight about the introduction of novel approaches for classifying electromagnetic data in geophysics.

CHAPTER II

EXPERIMENTAL CHARACTERIZATION OF CSEM RESPONSES FROM GEOLOGY AND CONDUCTIVE TARGETS*

1. Introduction

In near-surface geophysical applications, controlled-source electromagnetic (CSEM) methods are widely used because they are capable of detecting and mapping the distribution of fluids, large voids, faults or lithological contacts (Nobes, 1996; Tezkan, 1999; Pellerin, 2002; Everett and Meju, 2003).

The CSEM method also can detect and, to some extent, classify man-made targets such as pipes, underground structures and unexploded ordnance (UXO) (Chesney et al., 1984; Qian and Boerner, 1995; Huang and Won, 2003). The method is sensitive to subsurface spatial variations in electrical conductivity. While geological structures such as faults can generate recognizable CSEM signatures (Seaborne et al., 1979), the presence of natural structures with no clearly -defined dominant length scale gives rise to coherent but spatially rough CSEM responses that, until recently, have been ignored for inferring additional physical and geometrical properties.

A major challenge facing CSEM geophysicists involves detection and classification of man-made targets within signal-generated spatially-non-stationary geological noise. The target may be the object of investigation, as in unexploded ordnance (UXO) or geotechnical studies; or it may be a source of cultural noise (also called clutter, such as metallic debris, fences, abandoned cars, etc. that are not the object

*Most of the data reported in this chapter is reprinted with permission from “Target signal enhancement in near-surface controlled-source electromagnetic data” by A. Benavides and M. E. Everett, 2005. *Geophysics*, Vol. 70, G59-G67. Copyright 2005 by SEG.

of interest) that needs to be characterized so that correct subsurface interpretation of geology may be achieved. Advances have been reported in the detection of targets such as ships and icebergs using sea-surface-scattered radar signals along with a fractal description of the background sea clutter (Lo et al., 1993). The aforementioned approach is applicable to the CSEM method, keeping in mind the change in the governing physics from radar wave propagation to electromagnetic diffusion. Both problems contain roughly an equivalent level of complexity. The goal is to extend the techniques developed for radar to CSEM detection and buried target classification.

In regards to signal-generated geological noise, results already obtained (Everett and Weiss, 2002) from one-dimensional profiles show that a typical CSEM-response spatial profile has (i) a power-law spectral density $|A(\kappa)|^2 \sim \kappa^{-\beta}$ where κ is the wavenumber and (ii) the distribution of the differences between successive values is stationary. These two properties are typical of fractional Brownian motion (fBm), a self-affine fractal signal. The observed 1-D profiles are consistent with the superposition of a rough fBm background signal due to geology overlain by a much smoother signal due to man-made targets, which are inherently of a more regular geometric nature. The fBm component can not be explained in terms of instrument drift, survey procedures or external sources of incoherent noise (Benavides and Everett, 2002).

Self-affine data series have several special characteristics. The spectral exponent β is bounded ($1 < \beta < 3$); the phase is random; and the increments (the difference between two successive data values) corresponds to fractional Gaussian noise, (i.e. stationary signals with no long-range correlations). Finally, the variance may increase monotonically with series length and has power-law dependence (Turcotte and Huang, 1995).

In this study, the fractal nature of electromagnetic responses is further tested by the acquisition of additional profiles and two-dimensional (2-D) maps. This involves

analysis of apparent conductivity maps acquired using frequency domain CSEM systems. Maps of receiver voltage at a specific time gate would be appropriate in the case of time domain CSEM data.

We perform 1-D CSEM profile analyses using the Fourier Transform (FT), discrete wavelet transform (DWT) and variogram techniques. A standard Fourier analysis gives the spectral content of the signal, but provides no information regarding from where in space those spectral components originate. FT is not generally a suitable technique for non-stationary signals, although it can be used for non-stationary signals if we are interested in the spectral components existing in the signal, but not in where these occur. In the case of self-affine processes, FT can be used to determine power-law trends in the spectral content (Turcotte, 1992).

Wavelet analysis has been used in the analysis of certain types of geophysical time series, such as climate signals (Torrence and Compo, 1998). The idea of wavelet analysis is not new. It basically projects a signal $s(x)$ onto a set of localized basis functions, or wavelets. Multi-scale analysis is then performed using a set of these wavelets, each scaled by some amount a . The signal $s(x)$ is analyzed at each scale by translating the wavelet by some amount b . The coefficients

$$C(a, b) = \int_R s(x) \frac{1}{a} \Psi^* \left(\frac{x - b}{a} \right) dx \quad (2.1)$$

measure the similarity of the data to the basis wavelet. The function $\Psi[(x - b)/a]$ describes a wavelet $\Psi(x)$ scaled by $1/a$ and translated by b (Ψ^* denotes the conjugated of the complex function Ψ). Under conditions of low S/N ratios, a target response may be masked by the background signal. Using the wavelet transform, a change in the size of a given coefficient at the target location can be detected at different scales in the location-scale $(x - a)$ -space. Thus, wavelet analysis may reveal otherwise hidden signals if they are compactly located and possess a characteristic length scale.

Equation 2.1 can be used to analyze the relevant scales a given signal has. If we want to separate certain components without losing waveform information, DWT becomes a powerful tool. DWT analysis performs decomposition of the signal by extracting wavelet coefficients corresponding to large-scale (approximation coefficients A_j) and small-scale (detail coefficients D_j) features. This is done using a scaling function $\phi(x)$ and a mother wavelet $\psi(x)$ (Daubechies, 1992). The choice of the scales a and translations b is confined to the discrete set $a = 2^j$ and $b = k2^j = ka$ (for $j = 0, 1, \dots$). The general wavelets for multi-scale analysis are written as follows:

$$\Psi_{j,k} = \frac{1}{2^{j/2}} \Psi[s^{-j}(x - k)] \quad \phi_{j,k} = \frac{1}{2^{j/2}} \phi[2^{-j}(x - k)]. \quad (2.2)$$

Multi-scale analysis may be understood in terms of a decomposition tree where each branch corresponding to a scale j contains an approximation A_j and a detail D_j of the signal $s(x)$ that comprises small scale features of the signal at that scale. An approximation A_j is composed by the sum of all the previous details D_j ($j \leq J$). Then, the reconstruction of the signal $s(x)$ at any level j is achieved by adding the approximation A_j to the sum of all the previous details D_j . In the case of orthogonal wavelets, the approximations are orthogonal to the details and the details form an orthogonal basis (Mallat, 1998). Filters of different cutoff frequencies are used to analyze the signal at different scales. The signal is passed through a series of high-pass filters to analyze the details, and it is passed through a series of low pass filters to analyze the approximations. The same procedure is used to reconstruct the signal from the approximation and details with appropriate reconstruction, or inverse, filters.

Finally, variogram analysis is distinct from WT; it measures the variance $\gamma(h)$ of a series of length N as a function of the distance between pairs of data samples

$$\gamma(h) = \frac{1}{N} \sum_{i=1}^N [s(x_i) - s(x_i + h)]^2 \quad (2.3)$$

From the variogram, it is possible to estimate correlation lengths and infer the existence of fractal structures (Carr, 1995) even when the data are not acquired on a regular grid.

This paper describes results obtained by processing apparent conductivity profiles and maps using Geonics EM31 and EM34 controlled source instruments (McNeill, 1980) with and without the presence of highly conductive targets. We establish criteria to decide whether a target is indeed present and where it is located. We performed FT, wavelet transform algorithms and variogram analysis to assess which approach can perform the best and, at the same time, to validate the results using different data processing method. Additionally, we show how background analysis, target location, and geological noise rejection is achieved using the EM31 with the coil array oriented parallel (in-line) and perpendicular (broadside) to the profile direction. Our long-term goal is to develop algorithms for CSEM data that can be used to detect and classify buried, near-surface man-made conductive targets such as pipelines, borehole casings, steel drums, and UXO under real geological conditions.

2. Methods

Three different sites located at Texas A&M University (Riverside Campus, see Figure 3) 10 km west of College Station in Brazos County, Texas were selected for the designed experiments of this study. An additional location, in southeast Texas was chosen to test some of the performance of the signal processing under real conditions (see Figure 4).

The geology of the region consists of floodplain deposits of clayey sands, silts and clays, with high lateral continuity across the survey area. The sites are topographically flat, grass-covered, and without any discernible man-made electromagnetic noise

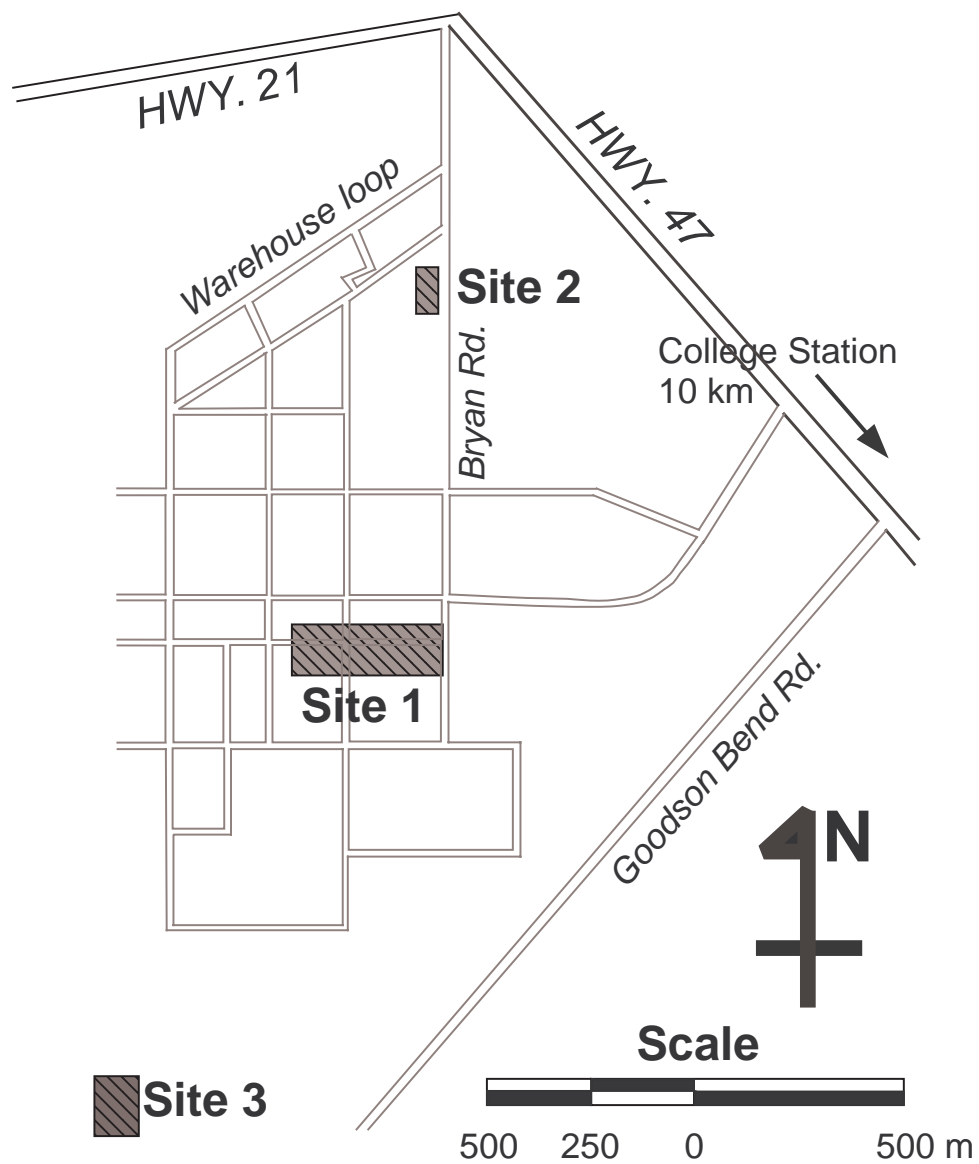


Fig. 3. Texas A&M Riverside Campus with the site positions used in the present work.

sources to distort CSEM observations. The Riverside campus was used as a military base during the 1940's so that it is possible, but not certain, that some metallic debris lies within the subsurface although no surface exposure can be visually detected. The land is mainly pastures with a few small buildings used for research, training, storage and administration purposes.

Both EM31 and EM34 systems were used (McNeill, 1980). They each consist of a transmitter-receiver (Tx-Rx) coplanar coil pair separated by distance s , the intercoil spacing, operating at a single frequency. The equipment in typical Brazos county soils operates at low induction numbers so that a linear relationship exists between the measured quadrature response of the subsurface and apparent conductivity σ_a (McNeill, 1980).

The EM31 has intercoil spacing $s=3.66$ m and, assuming a homogeneous half-space, approximately 60% of the induced signal comes from the uppermost 2 m. The EM34 operates at three different intercoil separations ($s=10, 20$, and 40 m) with the coils arranged horizontally (vertical dipole mode, VDM) or vertically (horizontal dipole mode, HDM) giving an estimated depth of investigation of $1.5s$ and $0.75s$, respectively (McNeill, 1990). For this study, the EM34 measurements were performed using VDM at $s=10$ m.

EM31 and EM34 profiles were conducted along a 152 m-long E-W transect at site 1 (Figure 3). The transect crosses near-surface features such as N-S gravel and pavement roads, shallow ditches and a buried drainage culvert. The station interval is $x=1.0$ m. The purpose of these profiles is to study how background geological noise spectral content is affected by the presence of target responses at the different depths of penetration afforded by the instruments.

Additional EM31 measurements were acquired at site 1 with the Tx-Rx array oriented parallel and perpendicular (P-array and T-array, respectively; see top of

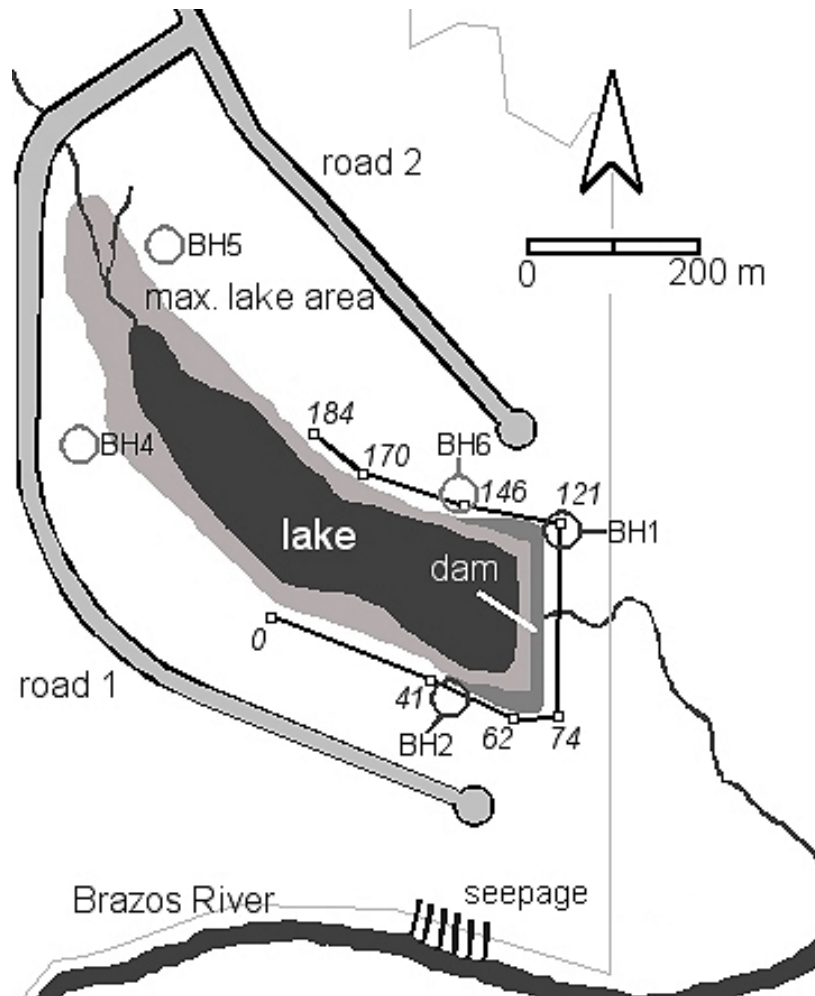


Fig. 4. General map of lake site in south Texas. The survey profile is shown as a polygon line surrounding the east half of the lake. The numbers in italic indicates the station number of each vertex (squares). Large gray circles indicate the only available information about borehole locations.

Figure 5) to a second E-W profile . The profile length is 64 m and the station interval is $x=0.5$ m. In order to generate target signals, conductive metal drill-pipe sections were laid out on the ground surface, oriented perpendicular to the profile. The EM31 profile was measured three times using different pipe configurations, as depicted in Figure 5 (middle). The apparent conductivity profiles were pre-processed to correct for drift due to battery power losses, the mean value was subtracted from the profile, following which FT, DWT and variogram signal analyses were performed.

Two experiments using the EM31 in P and T arrays were carried out in order to map the 2-D apparent conductivity distribution under two conditions: pure background noise; and in the presence of steel pipes. First, a survey was performed at site 3 (see Figure 3) using a regular 1 m grid in an area 40×23 m (920 measurements) in order to register a typical geologic noise response. Second, a survey was performed at site 2 using five pipes, arranged according to Figure 5(bottom). The station interval is $x=0.25$ m and the line spacing is 1 m in an area 30×6 m. A total of seven lines were surveyed, giving 847 apparent conductivity measurements per conductivity map.

The data from these two surveys were processed using a two-dimensional FT. The total spectral density $|A(\kappa)|^2$ was plotted as a function of the radial wavenumber k (where $k^2 = k_x^2 + k_y^2$). Additionally, log-log plots of the omni-directional variogram as a function of the lag distance h between data pairs were used to infer correlation length and the presence of fractal behavior. The latter can be recognized if the log-log variogram includes a linear trend, implying a power-law dependence on the lag distance. The slope m is related to the fractal dimension D by the expression:

$$D = \frac{F - m}{2} \quad (2.4)$$

where $F=2$ in the case of 1-D profiles (Carr, 1995) and $F=3$ in the case of surfaces (Herzfeld and Overbeck, 1999). In the wavenumber domain, fractal signals have

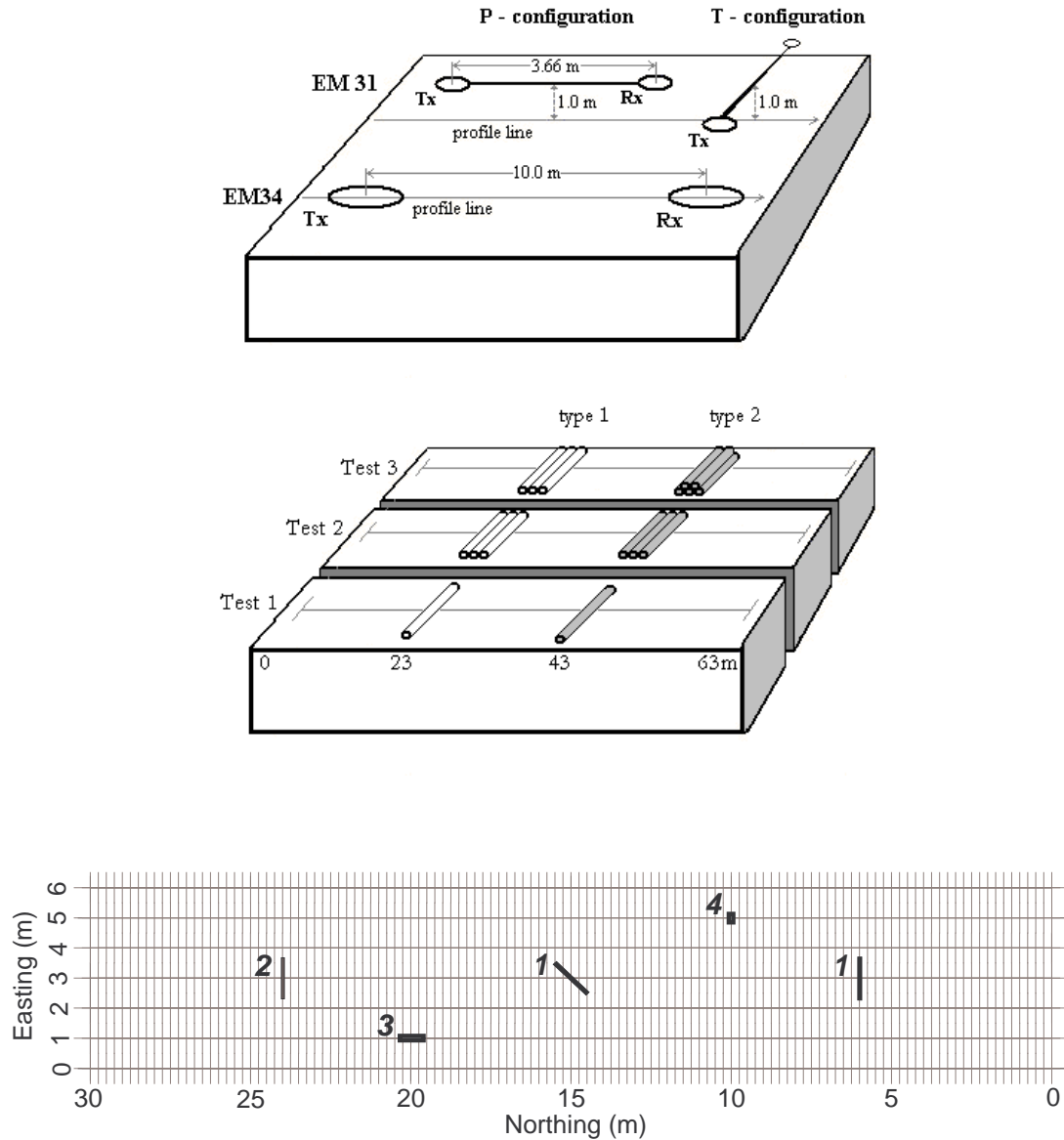


Fig. 5. EM31 Experimental setup. Top: sketch of the experimental setup showing how the CSEM sensors were oriented: P= coils parallel to the measurement profile; T= coils perpendicular to measurement profile. The objects in drawing are not to scale. Middle: configurations used in the 1-D profile tests at site 1. Key: type-1= 155 x 12.4 cm drill pipe, type-2= 150 x 6.4 cm drill pipe. The objects in drawing are not to scale. Bottom: 2-D survey layout at site 2 using five different pipes as targets. Horizontal lines are the line positions. Each node in the grid represents the location of both P and T array conductivity measurements. Key: #1= 155x12.4 cm type-1 drill pipe, #2= 150x6.4 cm type-2 drill pipe, #3= 75x7.5 cm galvanized iron water pipe, #4= 28x7.2 cm drill pipe. The pipe in the center of the grid is oriented at 45o with respect to grid lines.

a power spectrum with a power-law trend $|A(\kappa)|^2 \sim \kappa^{-\beta}$, where β is the spectral exponent. The fractal dimension D can be compared with β using (Turcotte, 1992)

$$D = \frac{G - \beta}{2}; \quad (2.5)$$

where $G=5$ for 1-D profiles, and $G=7$ for surfaces. In this way it is possible to compare the fractal dimension of a given signal using FT and variogram analysis.

Another processing alternative is to apply DWT to separate background from target signals. The final example comes from a multi-method, reconnaissance survey for assessing the existence of water seepage from an artificial lake (south-east Texas). A Geonics EM34 in vertical dipole mode configuration (intercoil separation of 10 m) was used. A map of the area and profile line is shown in Figure 4. The 920 meter-long profile consists of 185 apparent conductivity measurements using a station interval of 5 m. It turns out that around the lake there is a set of buried monitoring boreholes, poorly positioned in the available map (the borehole number sequence in the map does not have BH3). The purpose of the DWT was two-fold: i) to determine a possible high conductivity path that could link the artificial lake with some seepage observed on the nearby Brazos River, and ii) Try to locate the positions of three of the boreholes.

3. Results and discussion

A. Geological noise

A cursory inspection of the 6 different EM31 profiles in Figure 6 indicates that the CSEM signal including both targets and background has similar fluctuations for each type (P and T arrays) of surveys. These are the data from the 64 m-long profiles in which the various pipe configurations were laid out on the ground as in Figure 5

(middle). The data were gathered during a 7-hour period. Instrument response or environmental noise were discarded by repeating the profiles several times. Clearly, there is a coherent signal of spatially rough nature coming from the geology, as suggested by Everett and Weiss (2002).

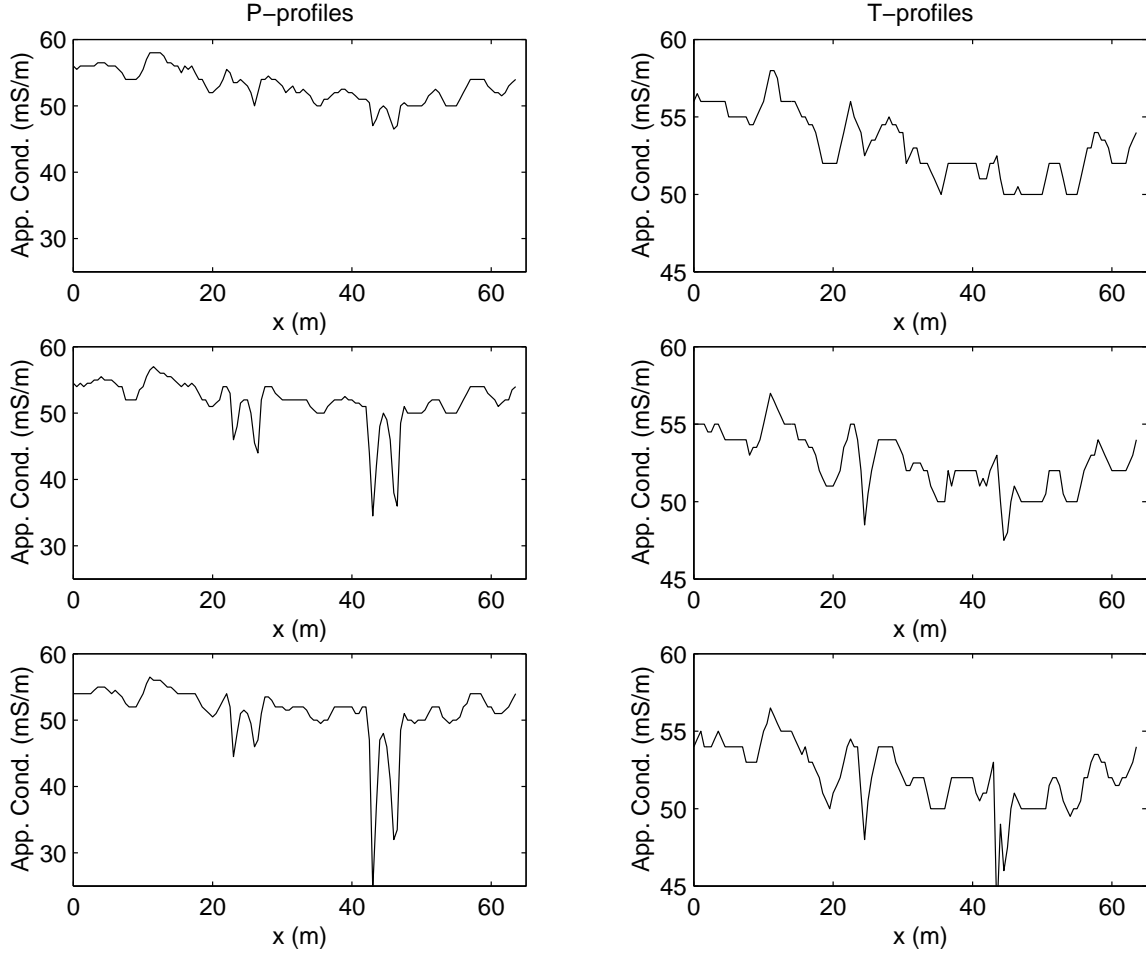


Fig. 6. EM31 apparent conductivity profiles from the pipe tests shown in Figure 5. Left column (from top to bottom, tests 1, 2, and 3) shows the P-array orientation apparent conductivity profiles. Right column (from top to bottom, tests 1, 2, and 3) shows the T-array orientation apparent conductivity profiles.

The amplitude of the target-response anomaly appears to be roughly proportional to the number of pipe segments laid out on the ground. The pipes are located at 23 m and 43 m along the profile. The target signature is bimodal and of high amplitude

in the P-arrays while the T-array target signatures are evident but more subtle. These anomaly morphologies are important for attempts at target classification using pattern recognition techniques. This concept is to be explored in later publications.

Figure 7 (top) shows data acquired along the 150 m-long profile at site 1 using both EM31 and EM34 systems. The anomaly detected at station 125 m is due to the buried drainage culvert. Notice that the two systems average different volumes of subsurface material and exhibit accordingly different wavenumber content in the background response. The fluctuations, which are as high as 10 mS/m in a 2 m interval, can not be explained in terms of hardware fluctuations since the equipment is properly calibrated, and as in Figure 3, the geological noise is highly reproducible if the survey is repeated carefully.

Figure 7(bottom right) shows the FT power spectrum for the EM31 profile. Analyzing the first part of the profile in Figure 7 (top) that contains just geological noise, the spectrum (irregular gray line) shows a power-law trend (straight gray line) with spectral coefficient $\beta=1.81$ ($D=1.57$), which corresponds to fBm. On the other hand, the spectrum (black line) of the entire profile shows energy leaking from intermediate wavenumbers ($\kappa \sim 3-9$ cycles/m) toward higher wavenumbers ($\kappa \sim 10-35$ cycles/m), with a steep decrease for the highest wavenumbers ($\kappa > 35$ cycles/m).

A similar result is obtained for the EM34 profile, but the spectral coefficient is $\beta=1.29$ ($D=1.86$). It may be shown theoretically that such a leakage of spectral power to higher wavenumbers is characteristic of a compact, deterministic signal overlying fBm. In a general sense, it is possible to infer the existence of a target if the spectral behavior is more complicated than the simple power-law exhibited by background geological noise. The target response in the wavenumber k domain affects a broad range of low wavenumbers (in this case, its influence is clear in the wavenumber interval 3-35 cycles/m) but the spectrum gives no information about location and

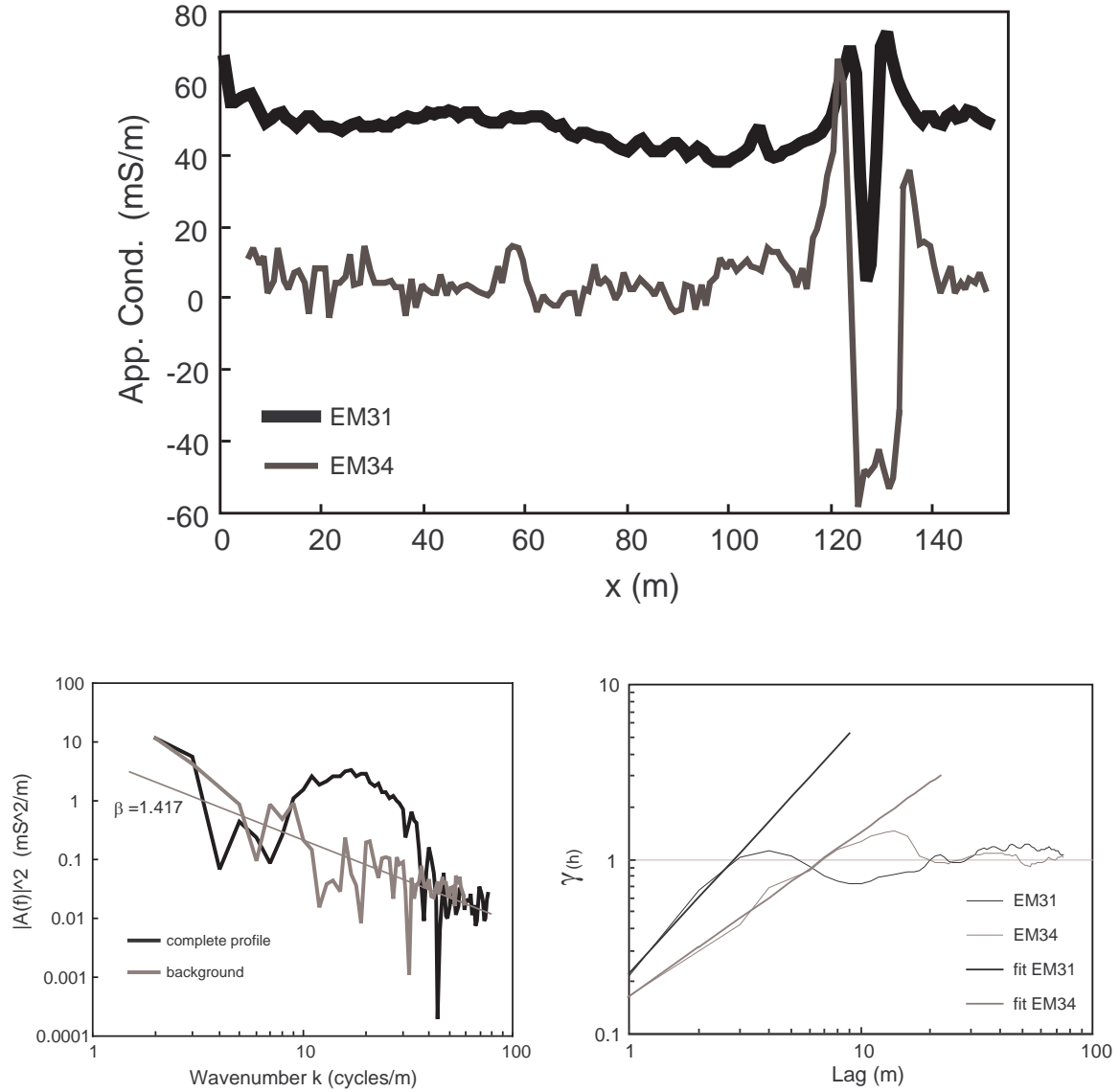


Fig. 7. Results from EM31 and EM34 apparent conductivity profiles for site 1. Top: EM31 and EM34 apparent conductivity profiles for site 1. Bottom left: Fourier Transform from the EM31 profile. The spectrum from the first 118 points (identify as background) includes only the response from the geology and presents a power-law trend with a least squares fit exponent $\beta=1.417$. The complete profile shows no power law trend. Bottom right: variograms for the EM31 and EM34 profiles. Straight lines are the least squares fit for the power-law portion of the variograms, giving fractal dimensions D of 1.28 and 1.53, respectively.

number of targets present.

The omni-directional variograms for both EM31 and EM34 profiles are presented in Figure 7(bottom right). Notice that the correlation length for the deeper-probing EM34 profile (approximately 7 m) is greater than that of the shallow-probing EM31 profile (approximately 3 m) and power-law trend is present for lags smaller than the correlation length (the fractal dimensions for the EM31 and EM34 using equation 4 are $D=1.28$ and $D=1.53$, respectively, corresponding to fBm).

Similar results are obtained if we process the EM31 profiles in Figure 6 using FT and variograms, thus providing information about the presence of targets. However, it is still critical to determine the location of such targets, especially for data with low S/N ratios. In this regard, we tested DWT to determine patterns in the wavelet coefficients ($x - a$ space) to see if we could filter out the geological noise and enhance target locations. The wavelet chosen was Coiflet1 (see waveform in Figure 8, top right), a compactly supported, orthogonal and symmetric wavelet (Daubechies, 1992) that resembles the general shape of the target responses.

If we want to determine the existence and approximated location of targets from the wavelet coefficients distribution, it is important to evaluate what the target response should look like in a noise-free profile. A typical type 1 pipe response (test 1, see Figure 5b) was obtained by averaging 10, short length profiles at different locations at sites 1 and 2. The scaling function and wavelet are given in Figure 8. Using DWT up to level two, we decompose the target response in one low-pass wavelet filtered profile and two high-pass wavelet filtered profiles at levels 1 and 2 (approximation A_2 plot and details D_1 and D_2 in Figure 8, respectively). The approximation A_2 reproduces fairly well the target response but, as we will show later, its frequency content is actually quite similar to the fBm geologic noise in the survey area. In the present case, more important target information is contained at the profiles D_1 and

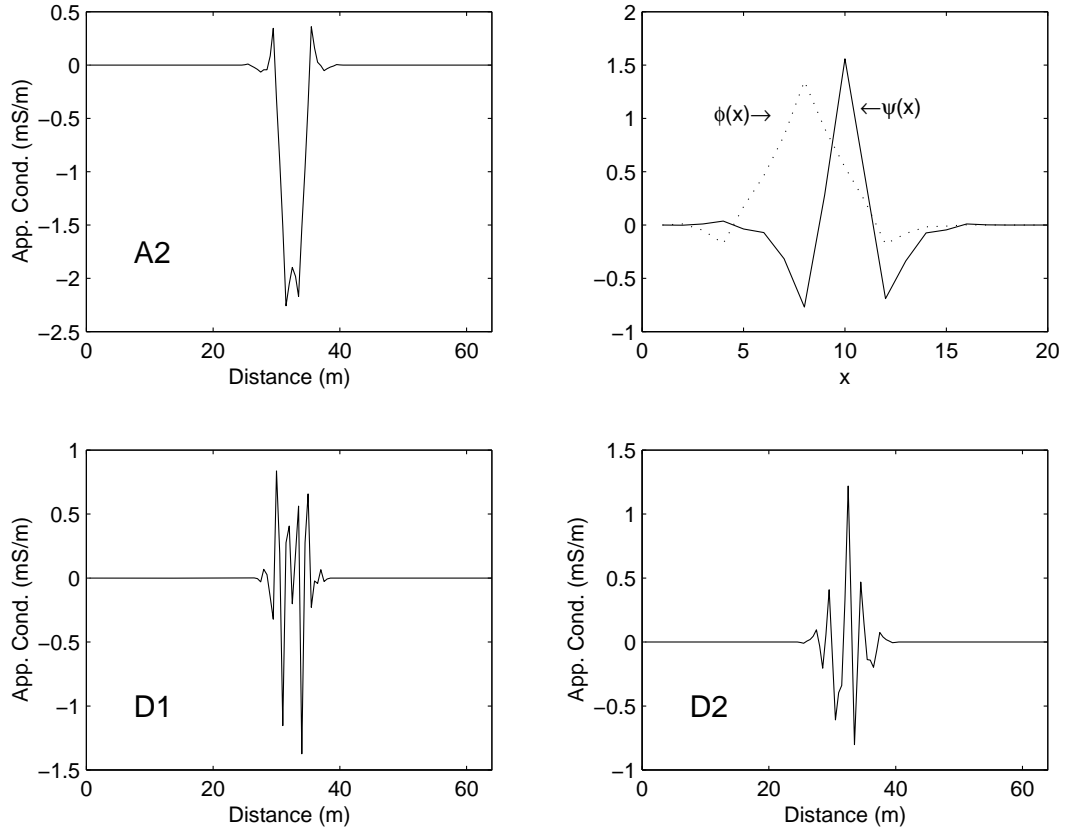


Fig. 8. Wavelet approximation and details for the average response of a type-1 pipe in parallel array configuration using DWT. Top left: the approximation (A_2) using the scaling $\phi(x)$ and mother $\Psi(x)$ Coiflet 1 wavelet (top right). Bottom: the details (D_1 and D_2).

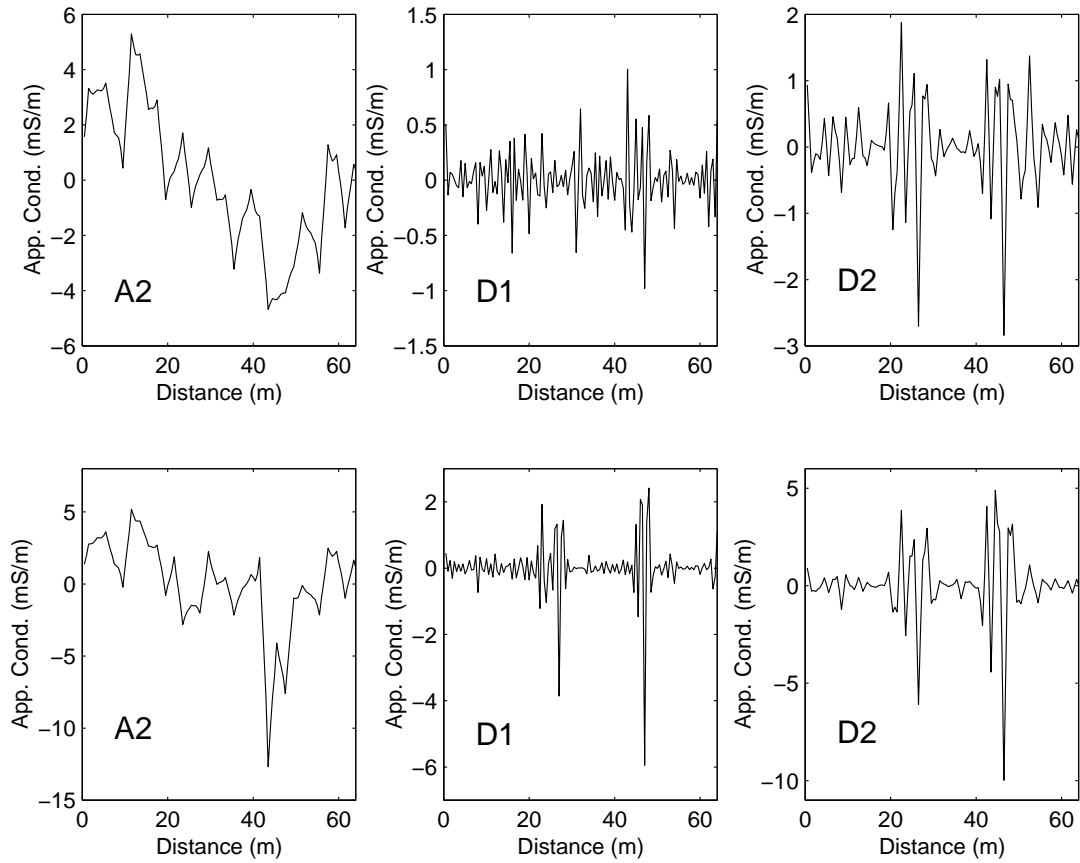


Fig. 9. Wavelet filtering of the EM31 profiles. Top: approximation and details for pipe test P1 using Coiflet 1 wavelet. From left to right: resulting reconstructed profile at approximation level 2 (A_2), details at levels 1 (D_1) and 2 (D_2). In this case, the contributions from the pipes are recovered at D_2 as high amplitude oscillations around the pipe locations (see Figure 5, middle). Bottom: approximation and details for pipe test P2 using Coiflet 1 wavelet. In this case, effects from the pipes are recovered in both approximations.

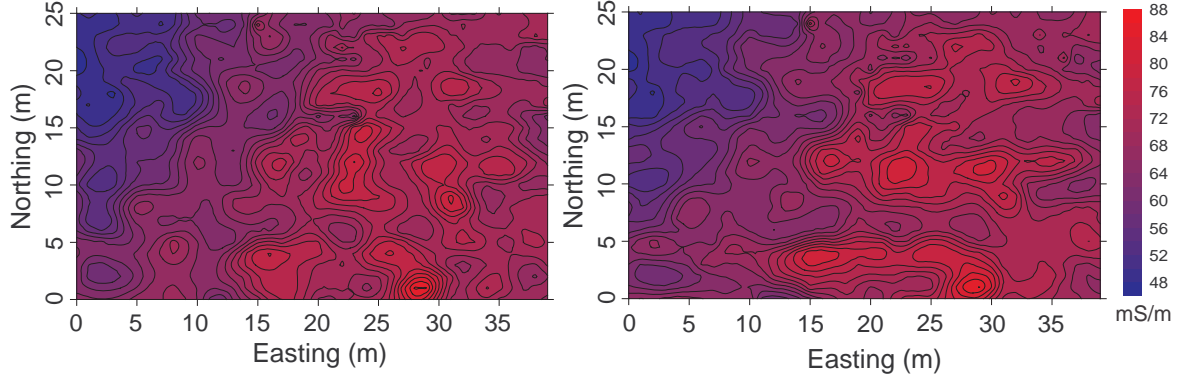


Fig. 10. Target-free EM31 apparent conductivity maps for the site 3. Left: P-array orientation (P). Right: T-array orientation (T). The number of stations in both maps is 920. Color scale in mS/m.

D_2 comprising low wavenumber signals.

Using the same Coiflet 1 wavelet, we process the actual data from P-array test 1 in the (top row, Figure 9). The DWT at the approximation A_2 keeps most of the geological noise along the profile while the detail profile D_2 maintains high amplitude oscillations at the pipe locations. We conclude that use of the DWT has been helpful to determine pipe locations in the case of low S/N ratio (i.e. when the range of apparent conductivity fluctuations is comparable or smaller than the amplitude of the target response).

A more favorable case (Figure 9, bottom row) shows the same plots of the previous figure but now applied to test 2 in the P-array orientation. Now, both pipes responses are identifiable the detail profiles D_1 and D_2 . The approximation A_2 preserves the low frequency components of the pipe response at location $x=43$ m, along with the bulk of the geological noise.

In the case of 2-D apparent conductivity maps, Figure 10 shows the EM31 apparent conductivity data obtained from site 3 (see Figure 3). There are no metal targets laid out at this site. Both P and T array maps show a regional trend of

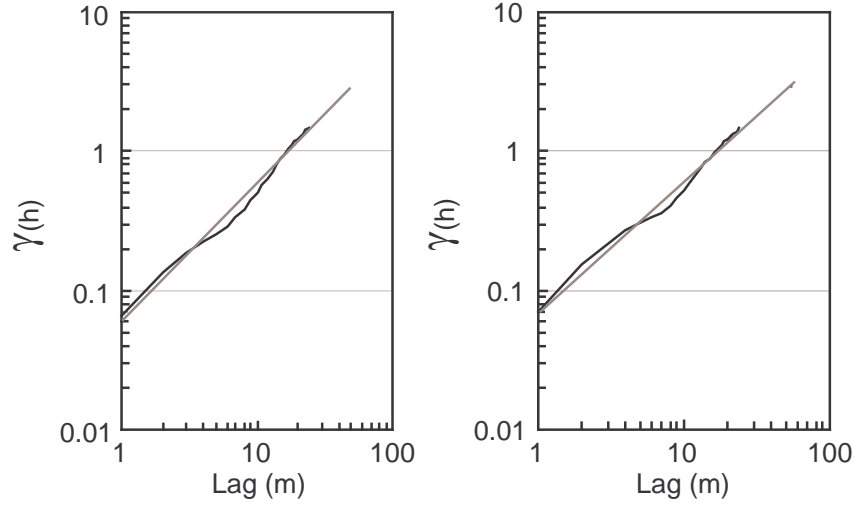


Fig. 11. 2-D omni directional log variograms from site 3 EM31 data (see Figure 10). Left: variogram for the P-array conductivity map. Right: variogram for the T-array conductivity map. Thick line represents the calculated variances. The thin straight line is the least squares fit.

decreasing conductivity toward the NW and a series of high and low conductivity variations at short length scales which are not related to the presence of any metal targets. The P and T array omni-directional variograms in Figure 11 show a striking power-law behavior for all lags with characteristic spectral exponents of $\beta=2.00$ and $\beta=1.94$ (fractal dimensions $D=2.50$ and $D=2.53$), respectively. The FT power spectra (Figure 12) show the same power-law trend for wavenumbers up to 13 cycles/m, beyond which the presence of the harmonics corresponding to the survey box size masks the background response. The spectral coefficients in both P and T array data correspond to fBm, with spectral exponents $\beta=2.82$ and $\beta=2.78$ (fractal dimensions $D=2.09$ and $D=2.11$), respectively.

The target-free omni-directional variograms in Figure 11 contrast with those resulting from the 2-D pipe experiment performed at site 2 (see apparent conductivity maps in Figure 13). Figure 14 shows that, in the 2-D case, instead of a monotonic

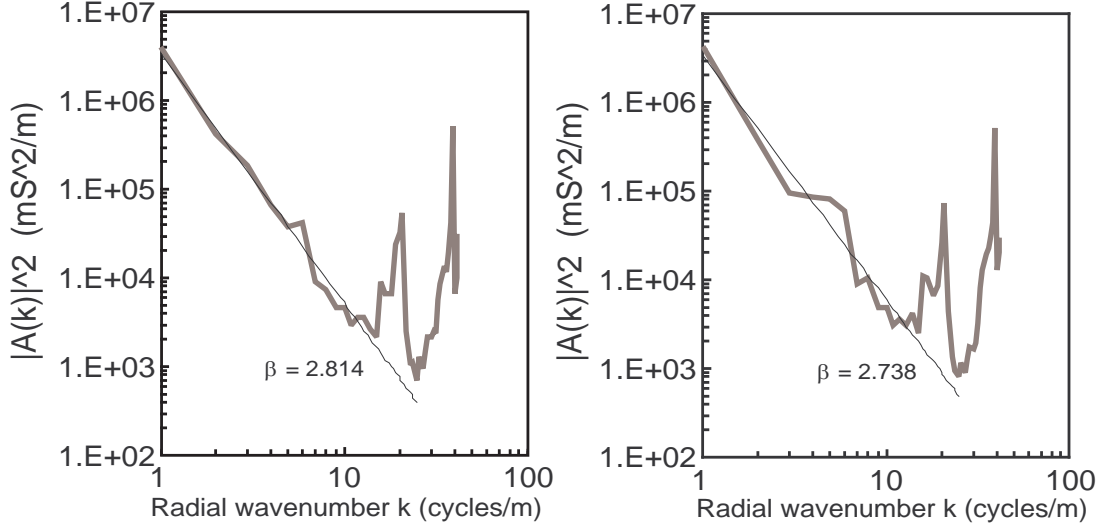


Fig. 12. 2-D FT for the P-array (left) and T-array (right) orientation conductivity maps. The abscissa variable corresponds to the radial wavenumber k . The spectrum follows a power-law trend, except for the peaks at high k which corresponds to the spatial dimension of the conductivity map. The least squares fit gives a spectral exponents $\beta=2.814$ ($D=2.08$) and $\beta=2.738$ ($D=2.1$), respectively.

power-law trend, the variogram present a correlation length around 5 m, which is roughly the distance between pipes. The result is consistent with the earlier results obtained in profiles, again supporting the interpretation of a non-fractal target signal superimposed on power-law geological noise.

B. Target-background separation

An approach to enhance target detection is to pre-process a CSEM response profile to take advantage of the fact that the background noise is spatially correlated. This is done to emphasize the signature of potential targets and to suppress geological noise. One approach is to combine P-array ($P(x)$) and T-array ($T(x)$) data into a single processed signal, $C(x)$. Figure 15 shows the function $C(x) = T(x) - P(x)$, for

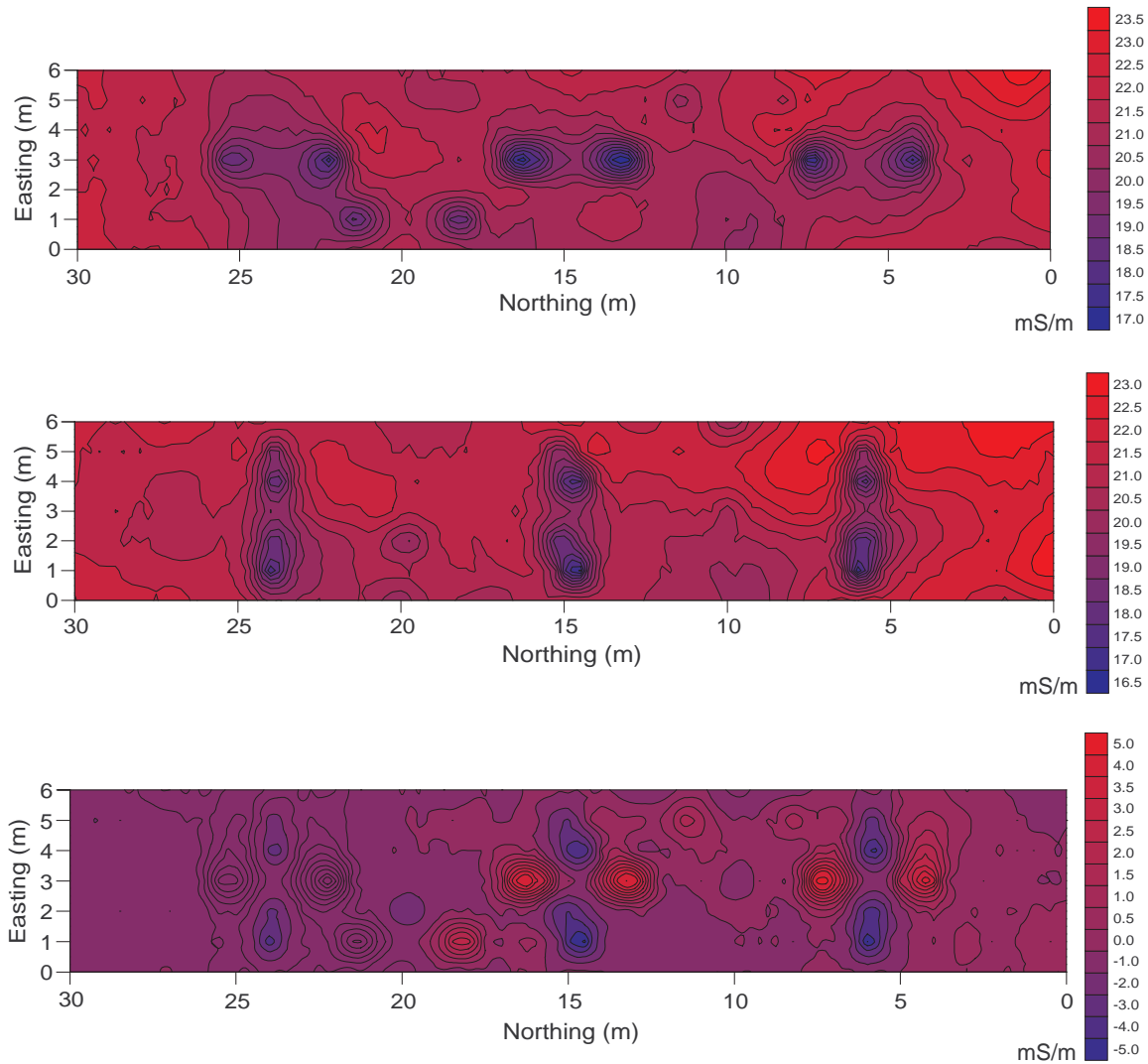


Fig. 13. EM31 apparent conductivity maps in the presence of pipes. P (top) and T (middle) array apparent conductivity maps using the layout shown in Figure 5. Target anomaly orientation depends on the array orientation relative to the target orientation; the characteristic anomaly length corresponds to the intercoil separation (s). Bottom: composite $C(x, y) = T(x, y) - P(x, y)$ conductivity map. Most of the background signal has been rejected, and the resulting target responses look like a quadrupolar anomaly. Even the small drill pipe at the position ($N=10, E=5$) has been enhanced. The shape of the 45° oriented pipe ($N=30, E=3$) appears distorted in that direction because of the contribution of the T-array conductivity values.

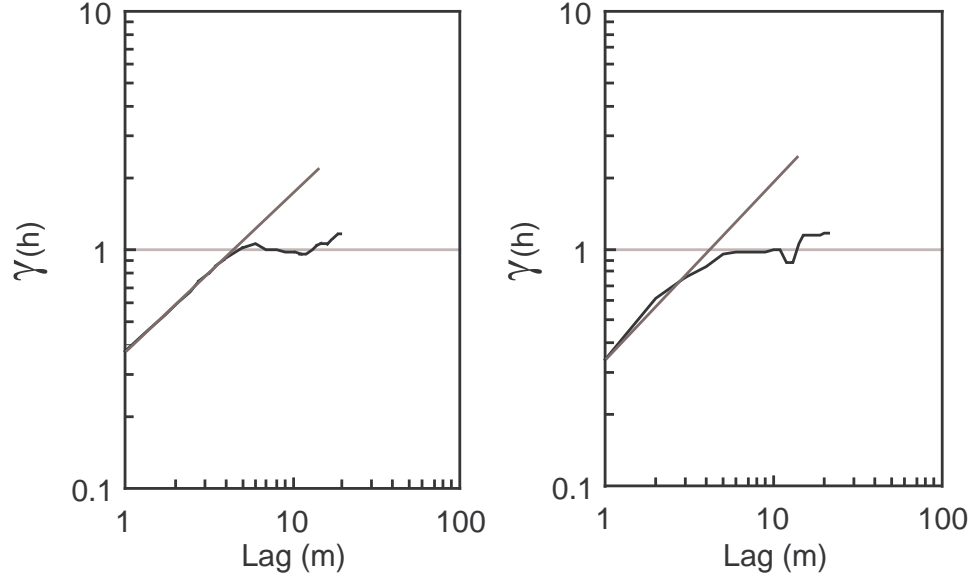


Fig. 14. Log-variograms for the P (left) and T (right) array apparent conductivities of Figure 13 (top, middle). The straight line corresponds to the least squares fit.

the P and T array pipe tests in Figure 3. The target signals are now clearly evident above the background. The target signature can be re-shaped into a single-maximum response using matched filters (Robinson, 1972) but this is beyond the scope of this paper.

The application of this background noise rejection approach using 2-D conductivity maps of site 2 is illustrated in Figure 13. Both P and T array conductivity maps (Figure 13; top, bottom) produce double peak anomalies for each pipe, arranged according to the TX-RX orientation. Notice that for elongated targets, EM31 P or T array maps does not yield much information but the composite map $C(x, y)$ does by showing elongation of the quadrupolar anomalies. The background response at distances far from the pipes is similar in both maps. A $C(x, y)$ conductivity map (Figure 13, bottom) shows that the pipes response becomes a quadrupole-like anomaly, centered on each pipe, while the background signal is greatly reduced.

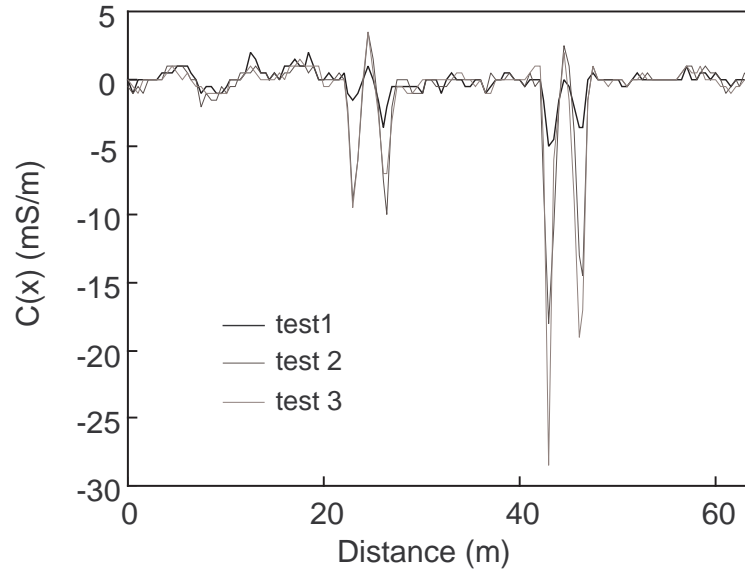


Fig. 15. Composite $C(x) = T(x) - P(x)$ for the pipe tests in site 1. See Figures 5(bottom), and 6) for details.

In order to test the previous finding in more realistic exploration conditions, one type-1 pipe was buried at site 3 to 0.15 m depth, transverse to the profile direction. Figure 16 shows the layout and the P and T array conductivity maps, along with the combined $C(x, y)$ conductivity map. There is almost no information about the presence of the pipe on the individual P and T array maps, but $C(x, y)$ shows the quadrupolar anomaly centered at the pipe location. Comparing with the results shown in Figure 13, it is clear that the combined effect of burial depth and conductive overburden has somewhat reduced the S/N ratio of target response.

The application of wavelet analysis to enhance borehole casing locations is next. Figure 17 shows the original and filtered profiles. The approximation A2 (identify as filtered profile) shows a relatively high conductivity zone in the station interval 35-65, just north of the observed seepage; it does not show a clear high-conductivity zone that can be associated with seepage but it is certainly the area with the highest

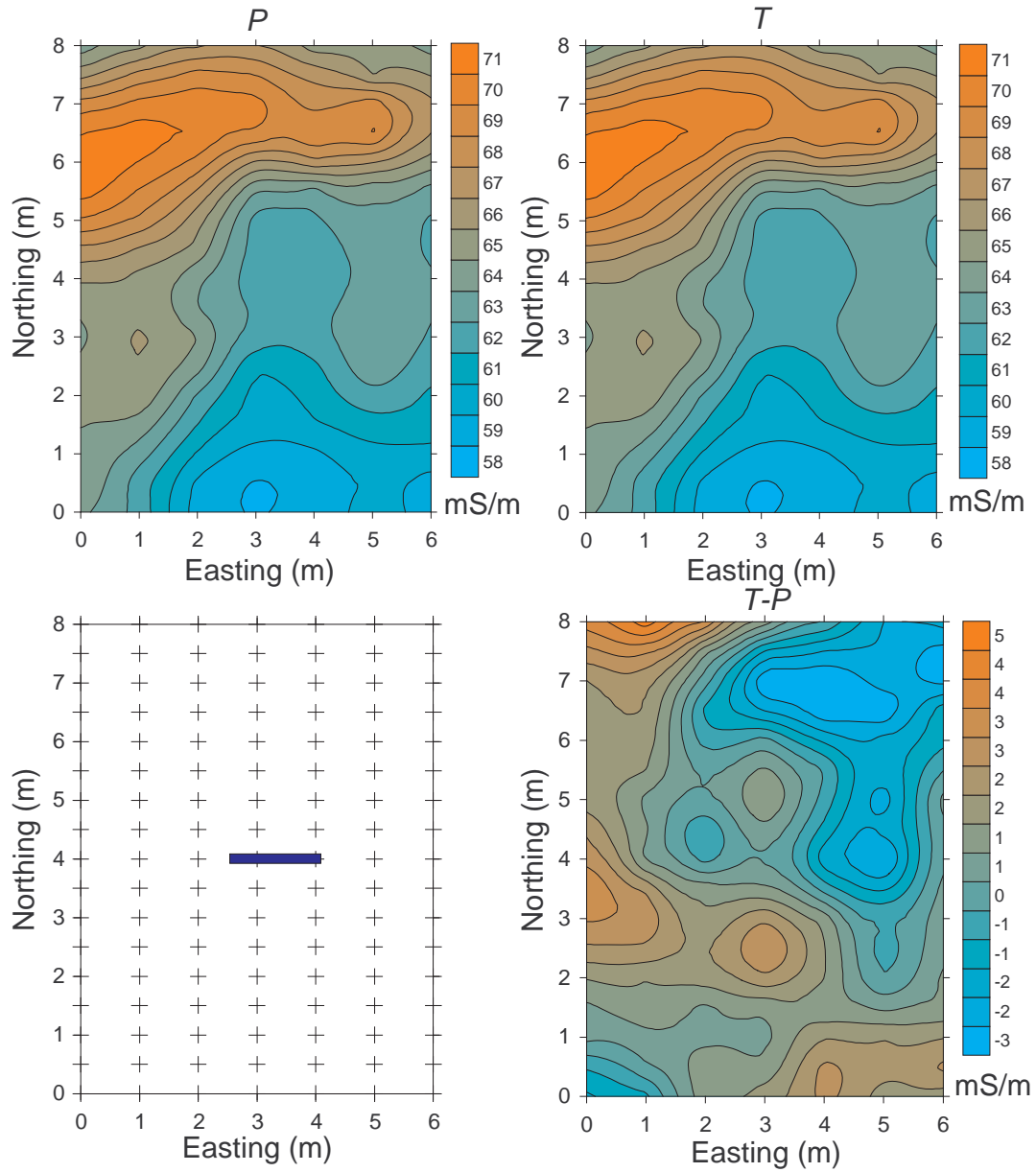


Fig. 16. Controlled EM31 buried pipe test. Top row: P and T array apparent conductivity maps obtained from the experimental setup show at bottom left. The $C(x, y)$ conductivity map (bottom right) reduced the background contribution enough to recover the quadrupolar-like response from the pipe. Notice the strong target amplitude reduction respect the previous tests and the increase of anomaly size.

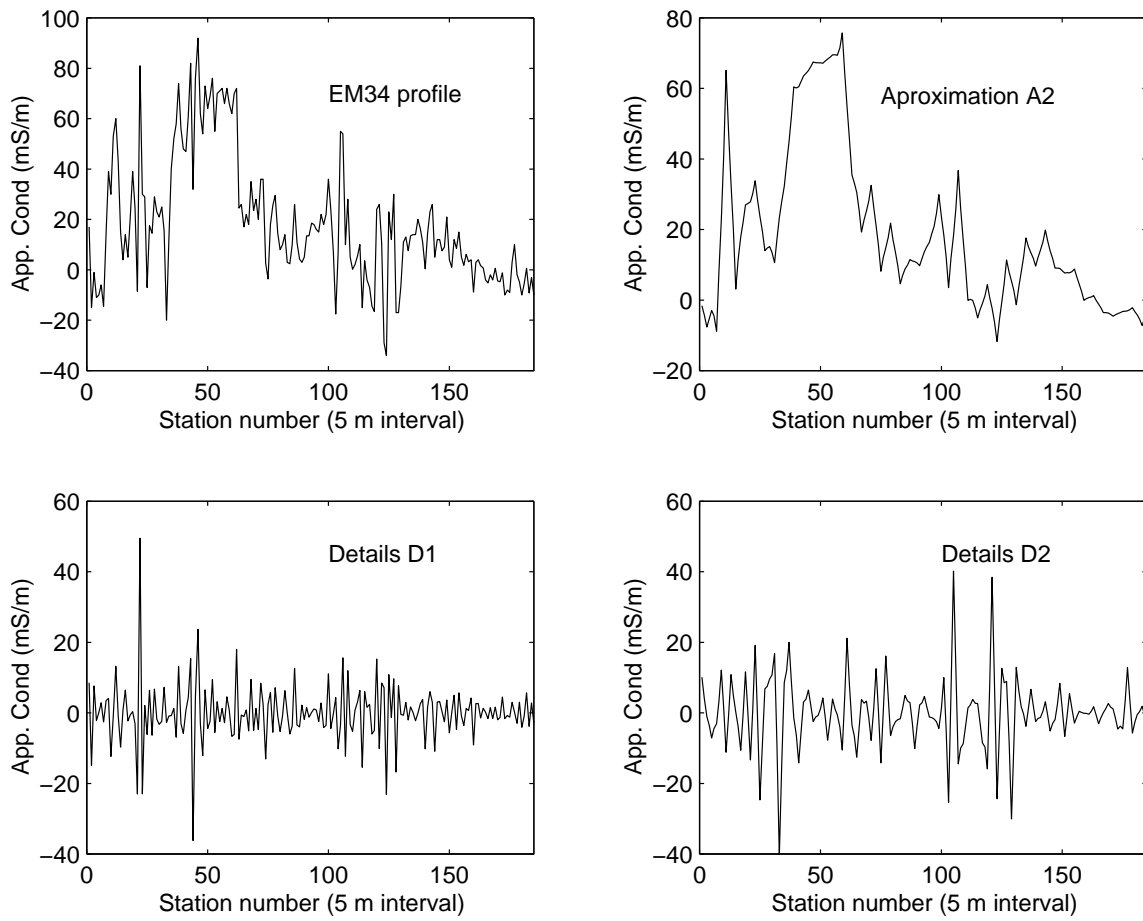


Fig. 17. EM34 profile data and processing results at the lake site using DWT. Top: apparent conductivity profile shows a superposition of a smooth signal affected by a high amplitude small scale noise. The next three profiles show the DWT approximation at level 2 (A_2) and details (D_1 and D_2) using Coiflet 1 wavelet.

apparent conductivity in the profile. Notice that the profile of details D2 has extracted the target signal corresponding to the dam's culvert at station 105, and the borehole BH1 at the NE corner of the dam. The profile with the details D1 extracted a target-like signature around station 44; possibly related with the borehole BH2. There is no EM34 target signature around the borehole BH6 location but another target-like response at station 22 (not associated with any available information) is clear on the detail D1 profile. Using this DWT processing, the resulting CSEM filtered profiles becomes easier to interpret by geotechnicians in terms of geology and man-made features.

CHAPTER III

INVERSION OF SPATIOTEMPORAL CSEM DATA USING A CONTINUATION METHOD*

Unexploded ordnance (UXO) detection, discrimination and removal represents a priority environmental and geotechnical problem in many regions of the world including Central America, Africa, and Asia. The common factor in cleanup operations is the huge final cost per UXO removed, mainly associated with unnecessary digging operations. The probability of UXO detection on documented test sites can exceed 90% in carefully executed geophysical surveys; however, despite recent advances in geophysical technologies and signal processing algorithms the false alarm rates remain unacceptably high.

1. Introduction

One of the challenges faced by geophysicists during cleanup operations is development of improved sensors and algorithms capable of recording target responses, isolating them from geological noise, and processing them in near-real-time without highly-trained operator intervention. A key step during data processing is model dimension reduction. The most widely-used dimension reduction method is inversion of the data using a low-dimensional forward model that sufficiently describes the target response. Under the requirements of UXO cleaning operations, the adopted inversion method must be fast, robust, and almost independent of the subjective bias and training of the instrument operator. The inversion should provide parameters whose values are

*Reprinted with permission from “Non-linear inversion of controlled source multi-receiver electromagnetic induction data for unexploded ordnance using a continuation method” by A. Benavides and M. E. Everett, 2007. *Journal of Applied Geophysics*, Vol. 61, 243-253. Copyright 2007 by Elsevier B.V.

invariant with respect to the sensor-target geometry but highly dependent on the shape and physical properties of the target.

A. The inverse problem

Inversion of geophysical data consists of finding a model vector \mathbf{b} that reproduces the most important features of a set of observations \mathbf{d} using a physically justifiable forward response $f(\mathbf{b})$. The forward response in most cases is non-linear. The "best" model \mathbf{b}_* minimizes a scalar cost, or objective, function that measures the goodness of fit between the observations and the forward response. The most widely used cost function is based on the L_2 least-squares norm. The non-linear inverse problem is traditionally solved by Newton-Raphson, Gauss-Newton, Levenberg-Marquardt (LM), Powell's methods (Madsen et al., 2004), or conjugate gradients (Nocedal and Wright, 1999) to cite just a few. Unfortunately, geophysical inverse problems are ill-posed and the methods to solve them oftentimes are unstable (Vasco, 1998) due to intrinsic non-uniqueness, over-parameterization, and/or large residuals. The latter is relevant for data with a large dynamic range, such as time-domain EMI responses. If the initial model vector \mathbf{b}_0 is far from the convergence region of a local or global minimum of the cost function, then oftentimes no solution with a tolerable misfit can be found.

B. Continuation methods

Continuation/homotopy algorithms have long been used for solving non-linear systems of algebraic equations. These methods, which are convergent for almost all choices of starting point, are broadly applicable to a wide range of non-linear problems (Watson, 1989; Rheinboldt, 1980).

In geophysics, continuation methods have been applied to the inversion of seismic ray-tracing data (Keller and Perozzi, 1983), to seismic travel-time tomography

(Vasco, 1994), and to inversion of multi-frequency EMI data (Jegen et al., 2001). Furthermore, the solution of the regularized inverse problem for seismic cross-well tomography (Bube and Langan, 1994), and seismic first-arrivals estimation for near-surface velocity structure (Vasco, 1998) has shown the advantages of continuation methods to evaluate the misfit vs. model roughness trade-off. An early use of the homotopy method based on formulating the EMI inverse problem as an equivalent set of polynomial equations is provided by Everett (1996).

C. Outline

A continuation method for inverting time-domain multi-receiver EMI data is presented. The method is based on a physics-based UXO forward response described by a stretched-exponential decaying magnetic spheroid. Near real-time processing rules out more accurate but time consuming forward modeling approaches using finite difference or finite element methods. The use of closed-form solutions for the EMI response of simple geometrical shapes in free space such as a hollow sphere (Mrozynsky, 1998) or a prolate spheroid (Ao et al., 2002) is not recommended for near-real-time data processing since these solutions are not general enough to discriminate the wide range of target shapes that could be encountered in a UXO geophysical survey. The use of the magnetic response of a spheroid or of a simple dipole response (Pasion and Oldenburg, 2001; Zhang et al., 2003) satisfies the near real-time requirement while providing target discrimination capabilities.

A longer-term goal is to develop new UXO discrimination capabilities based on time-domain EMI measurements utilizing multi-transmitter (Tx) multi-receiver (Rx) data acquisition. This paper, aimed as a step toward achieving this goal, is organized as follows. First, the continuation algorithm is presented along with a mathematical description of the forward response and the experimental setup. Second, a description

of the transient multi-receiver (i.e. spatiotemporal) EMI target response is accompanied by several inversion examples using both noisy synthetic data and the actual measured response of different buried ordnance.

2. Methods

A. Inversion

In this section we summarize briefly the principal aspects of the continuation method considered here. The data vector $\mathbf{d} = (d_1, d_2, \dots, d_n)^T$ represents the spatiotemporal EMI response from a buried target. Each datum is associated with a vector \mathbf{x} of independent variables, consisting of Rx coil location and orientation, Tx location and current, and a list of the time gates at which the EMI response is sampled. The physics-based forward response capable of predicting the data vector is defined as $f(\mathbf{x}, \mathbf{b})$, where $\mathbf{b} = (b_1, b_2, \dots, b_p)^T$ is the vector of model parameters. An auxiliary continuation parameter denoted as λ is introduced. For the inversion process, the $(p + 1)$ -dimensional search space is the one that is spanned by the augmented model vector (\mathbf{b}, λ) . The augmented cost function to be minimized is

$$S^\lambda(\mathbf{b}) = \frac{1}{2} \|F^\lambda(\mathbf{b})\|_2^2 + \sum_{i=1}^p G(\mathbf{b}_i). \quad (3.1)$$

Equation (3.1) consists of an augmented L_2 -norm misfit and a set of penalty terms G operating on the model vector \mathbf{b} . The standard and augmented residuals are defined as follows

$$F_i(\mathbf{b}) = f(\mathbf{x}_i, \mathbf{b}) - d_i, \quad (3.2)$$

$$F^\lambda(\mathbf{b}) = F(\mathbf{b}, \lambda) = F(\mathbf{b}) + (\lambda - 1)F(\mathbf{b}_0). \quad (3.3)$$

The augmented residual, equation (3.3), is a linear combination of the standard residual vector $F(\mathbf{b})$ and a constant vector $F(\mathbf{b}_0)$ describing the residual produced by the starting model \mathbf{b}_0 . The value of the continuation parameter λ determines the amount of change in the augmented residual with respect to the standard residual, equation (3.2). An initial augmented model vector $(\mathbf{b}_0, 0)$, selected a priori, forces the cost function to be zero. Small subsequent changes in any of the augmented model parameter lead to small residual departures from zero.

The purpose of the penalty terms in the cost function is to keep the model parameters within well-defined bounds. Out-of-bounds model parameters are penalized by a high cost. We have chosen a sigmoidal well penalty function characterized by a nearly constant zero cost inside the allowed region and steep walls rising to a very high cost at the boundaries

$$G(b_j) = A [2 + \tanh(\theta b_2) - \tanh(\theta b_1)], \quad (3.4)$$

$$b_2 = \frac{b_j - b_j^{max}}{b_j^{max} - b_j^{min}} \quad (3.5)$$

$$b_1 = \frac{b_j - b_j^{min}}{b_j^{max} - b_j^{min}} \quad (3.6)$$

In these equations, the adjustable parameter θ controls the steepness of the boundary wall and b_j^{max} , b_j^{min} are the pre-determined upper, lower limits for the model parameter b_j . Suitable amplitude A is determined after the standard misfit is calculated following the first iteration.

The least-squares problem is solved by tracking any possible $(p + 1)$ -dimensional path Γ that maintains the cost function near zero, $S^\lambda(\mathbf{b}) \sim 0$. The solution model vector \mathbf{b}_* is found at the intersection of the path Γ with the plane $\lambda = 0$ (see Figure

18).

The path-tracking through the augmented model space in Figure 18 can be accomplished in several ways, each providing the discrete equivalent of a smooth, continuous, non-bifurcating, non-divergent path (Garcia and Zangwill, 1981; Watson, 1989). In this work, path-tracking is based on a set of predictor-corrector steps. The predictor step consists of varying the augmented model vector (\mathbf{b}, λ) by $\Delta\lambda$ along the λ -direction. The corrector step is a Gauss-Newton (GN) sequence of iterations. At the k -th predictor-corrector step, the predicted model vector is $(\mathbf{b}_k, k\Delta\lambda)$. The corrector step at the $(l+1)$ -th GN iteration updates the model vector using (Villiers and Glasser, 1981)

$$\mathbf{b}_k^{l+1} \cong \mathbf{b}_k^l - \mu [\nabla^2 S^k(\mathbf{b}_k^l)]^{-1} \nabla S^k(\mathbf{b}_k^l). \quad (3.7)$$

The second term on the right-hand side of Equation (3.7) contains the negative gradient, and the Hessian matrix (inside the square brackets) scaled by stepsize μ . The gradient is calculated analytically while the Hessian is calculated using an approximation for small residuals (Nocedal and Wright, 1999) or forward differences. A singular value decomposition (SVD) algorithm is used to calculate the Hessian matrix inverse or its pseudo-inverse by using only the eigenvalues larger than some predetermined threshold value. The predictor step is controlled such that it is small enough to allow the GN corrector to converge in a few iterations.

Three corrector-step stopping criteria are used based, respectively, on i) variation of the updated model vector; ii) variation of the cost function value, and iii) the value of the gradient. Once the corrector step is terminated, the next predictor step is initiated. Following the zero path $S^\lambda(\mathbf{b}) = S_k(\mathbf{b}_k) \sim 0$ by this series of predictor and corrector steps from $\lambda = 0$ to its endpoint $\lambda = 1$ at high accuracy generally is not necessary. However, near the endpoint of the path, the predictor step-length and

corrector stop conditions must be tightened in order to maintain the accuracy of the path tracking and producing a more reliable final model vector.

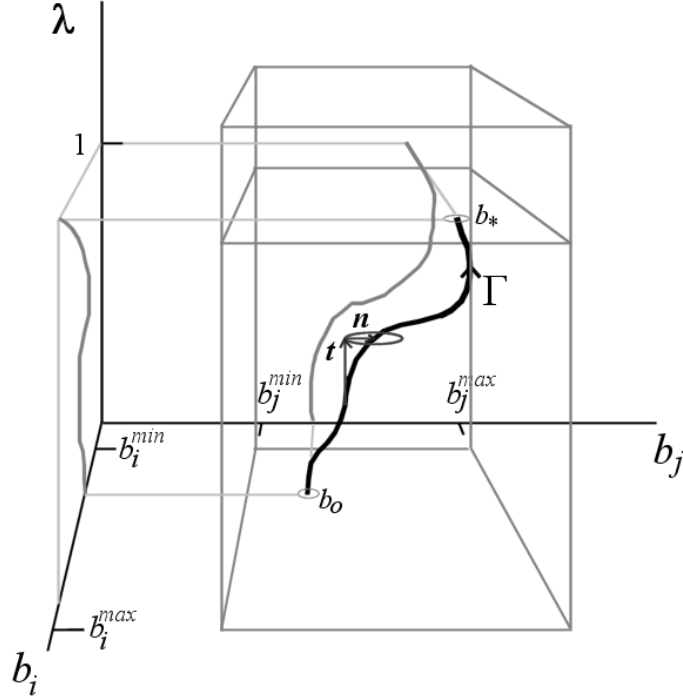


Fig. 18. Schematic of a continuation path Γ in a bounded model parameter search space (dimension 3). The path starts at the point $(\mathbf{b}_0, 0)$ following a series of predictor (t) step and corrector (n) steps. Every predictor point should be inside the convergence radius of the corrector in order to find the new continuation point. At the intersection of the path Γ with the plane $\lambda=1$ a solution of the least squares problem \mathbf{b}_* is found. Path projections on the planes $b_k - \lambda$ (thick grey curves) provide a way to monitor the variation of each model parameter during the inversion.

B. Forward response

The physics-based EMI forward response corresponds to a transient, decaying magnetic dipole source $\mathbf{m}(t)$ located at \mathbf{r}' that induces a voltage in a receive coil located at \mathbf{r} with normal vector \mathbf{n} of the coil axis. The spatiotemporal EMI response is

$$f(\mathbf{r}, \mathbf{r}', \mathbf{b}, t) = -\frac{\partial}{\partial t} \mathbf{T} \mathbf{m}(t) \cdot \mathbf{n}^{rec}, \quad (3.8)$$

where $T = \nabla \nabla (1/r)$. In order to model the response of axi-symmetric targets, we use an empirical stretched-exponential model for $\mathbf{m}(t)$ which describes the superposition of two magnetic moment vectors, one directed along the principal axis of the spheroid $\hat{\mathbf{a}}$ (subscript L), and the other pointing in the direction of the primary field \mathbf{B}_{tr} (subscript T)

$$\mathbf{m}(t) = (\gamma_L - \gamma_T) (\hat{\mathbf{a}} \cdot \mathbf{B}^{tr}) T_L(t) \hat{\mathbf{a}} + \gamma_T B^{tr} T_T(t) \hat{\mathbf{B}}^{tr}. \quad (3.9)$$

The polarizability parallel and transverse to the principal axis are denoted by γ_L and γ_T , respectively. Notice that each term on the right hand side of equation (3.9) decays at a different rate, $T_L(t)$ or $T_T(t)$, where

$$T_L(t) = -\frac{\partial}{\partial t} [t^{-\beta_L} \exp(-(s_L t)^{1-\beta_L})], \quad (3.10)$$

$$T_T(t) = -\frac{\partial}{\partial t} [t^{-\beta_T} \exp(-(s_T t)^{1-\beta_T})]. \quad (3.11)$$

In these equations, $\beta_{L,T}$ and $s_{L,T}$ are the stretching coefficients and decay rates, respectively. We choose the stretched exponential induced current decay for $T(t)$ because it provides a simple and effective temporal description of the EMI transient that agrees with our field experience. The stretched exponential form is physically justifiable since it describes an eddy current generation mechanism in a heterogeneous target that is based on a multiplicative process (Laherrère and Sornette, 1998). Similar multiplicative processes have been reported for remanent magnetization decay in spin-glasses and dielectric relaxation (Chamberlin et al., 1984). Stretched exponential decays are broad-band, continuous functions that can be interpreted in terms of an assemblage of interacting eddy current modes co-existing in a geometrically complex target.

C. Data acquisition

Controlled source EMI data were acquired using a modified Geonics EM63 time-domain configuration (Geonics, 2002). The commercial EM63 system uses a TX bipolar square waveform with base frequency of 5 Hz, delivering a maximum current intensity of 16 A. Transients are recorded at 26 geometrically time-spaced gates from $177 \mu\text{s}$ to 25 ms (see Appendix A). The EM63 is a cart-mounted central-loop configuration system. When it is used in wide-area surveying mode, its designed purpose, a few transients are recorded and assigned to the same station number. The distance covered between stations at a typical walking pace is about 0.1 m.

We modified the system geometry by separating the RX and TX coils (Everett et al., 2005). A custom-adapted 4.0 m cable supplied by Geonics and a simple 3.0 m long wooden structure allows multi-Rx location data acquisition while the Tx coil remains fixed, in this case, laid on the ground surface (Figure 19). The RX unit is moved along the wooden structure, sampling the temporal EMI response each 0.1 m along a 2.5 m profile in the x-direction. Raw target transients are stacked at each RX location to enhance the signal-to-noise ratio (SNR). Combining the temporal response at each station generates the full EMI spatiotemporal response of the target and background. Each spatiotemporal response corresponds to 25 RX locations and comprises about 1,500 raw transients.

We measured noise levels by analyzing the same amount of transients used for producing the standard datasets. It was found that the mean system noise floor is $20 \pm 70 \mu\text{V}$. The SNR of spatiotemporal data in the presence of a target is defined as the ratio of the maximum receiver voltage V_{max} over the sum of the average and standard deviation of the noise floor (VNF and NF, respectively): $SNR(dB) = 20 \log_{10}(V^{max}(mV)/(VNF + NF))$. In the case of the synthetic data

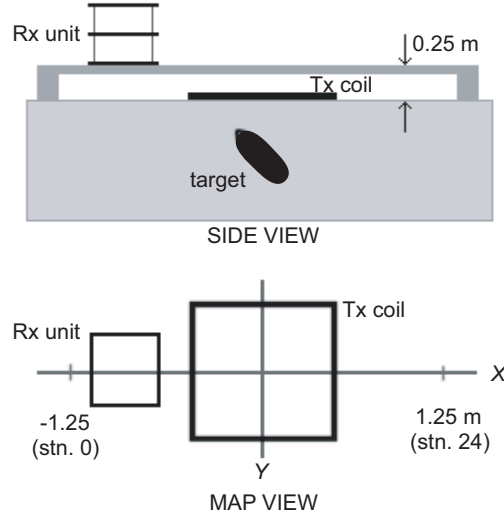


Fig. 19. Schematic of the experimental multi-receiver data acquisition setup using a modified Geonics EM63. The RX unit is moved along the wooden structure, 0.25 m above the ground surface. The detected target is positioned about the center of the Tx coil.

example, Gaussian noise (with a standard deviation of $0.2V_{NF}$) was added to produce a noisy version with $SNR = 20$ dB. This noise level is in some cases one order of magnitude larger than the one we found in EM63 transients of buried ordnance.

Spatiotemporal responses were acquired by burying a target at some depth (between 0.1 m and 0.5 m) centered on the TX coil. The central burial is not necessary but is chosen here since it achieves maximum TX-target electromagnetic coupling. Background responses were also measured before and after each acquisition protocol to enable an estimate of the background-subtracted free-space target response and to monitor TX current stability. For the present paper, we assume that the mutual electromagnetic coupling between the target and the background geological medium is negligible so that the total EMI response is a simple, linear superposition of both target and background responses. The background-subtracted (BGS) data (Figure 20) is then used during the inversion process. During actual UXO prospecting, the background response without the target present is not known. In that case, we recom-

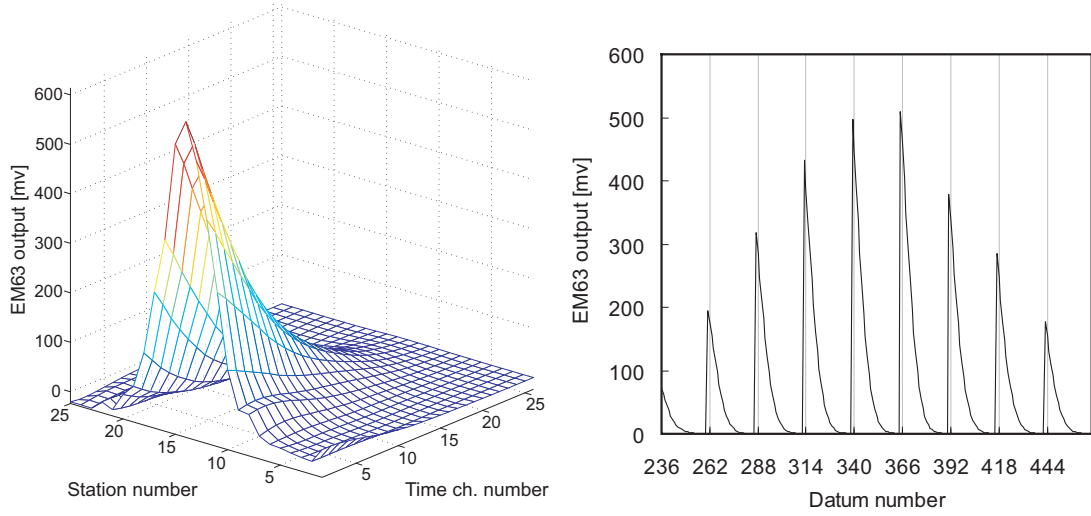


Fig. 20. Spatiotemporal data representation. Left: 50mm M22 ordnance spatiotemporal response as a surface in 3-dimensions. At each station, an averaged transient is recorded (lines running parallel to the time channel number axis). Each node in the spatiotemporal surface corresponds to a RX output voltage, or datum. Right: on the rest of the paper, the spatiotemporal response is represented as a 1-D vector of appended transients starting from station 1 to 25. Each transient consists of 26 data datums.

mend that BGS data are formed using spatial averaging of background EMI readings taken in the vicinity of a putative UXO target. Initial values for the model parameters \hat{a} and r_o were kept fixed for different model vectors chosen. The experimental setup constrains the horizontal location of the target while the vertical coordinate can be estimated by cursory inspection of the maximum, asymmetry, and width of the spatial EMI response.

3. Results and discussion

Four inversion examples are discussed here to illustrate the performance of the continuation algorithm. The 12-dimensional model vector in both cases corresponds to the forward model described by equations (3.8 through 3.11), and it is ordered as follows $\mathbf{b} = (\gamma_L, \gamma_T, \hat{a}, r_o, \beta_L, \beta_T, s_L, s_T)^T$. The decay rates s_i are expressed in 1/ms.

The initial model \mathbf{b}_0 is the same in all examples.

A. Synthetic data

For this first example, a synthetic data vector was generated using the following model vector $\mathbf{b}=(1.14, 0.08, 0.8551, 0.0, 0.5184, 0.12, 0.1, 0.25, 0.1, 0.3, 0.5, 3.5)^T$, corresponding to the target principal axis $\hat{\mathbf{a}}$ inclined 31.230° in the x-z plane. The target is located at burial depth 0.25 m below the $z = 0$ plane, and the longest eddy current decay is along the $\hat{\mathbf{a}}$ direction. The synthetic data were contaminated by Gaussian noise to produce a $SNR = 20$ dB. The synthetic TX coil consisted of a 1.0 m x 1.0 m square coil driven by 1 A current placed 2.5 cm above the ground surface to simulate our modified EM63 multi-RX data acquisition.

The continuation algorithm was initiated using the starting model $\mathbf{b}_0=(0.1, 0.1, 0.7071, 0.0, 0.7071, 0.0, 0.0, 0.3, 0.001, 0.001, 1.0, 2.0)^T$. The path step length was selected as $\Delta\lambda=0.05$. For the corrector step, stopping tolerances were set to 10^{-3} for the cost function and model vector variations, respectively. The maximum number of corrector iterations, at each predictor step, was set to 5 and a SVD threshold of 10^{-5} was established for the Hessian pseudo-inverse computation. Figure 21(top left) shows the exponential convergence of the inversion with predictor iteration number. The cost function value is reduced by almost three orders of magnitude. The path tracking in the augmented model space is visualized in Figure 21(top right). Smooth evolution of each model parameter is observed after the third predictor step. Variations of more than two orders of magnitude in transversal stretched coefficient β_T and decay rate s_T reflects the attempts of the algorithm to fit the model vector to the small amplitude decay in the direction of the primary field. The model vector found for $\lambda=1.0$ was $\mathbf{b}_*=(1.04, 0.08, 0.837, -9.68\times 10^{-12}, 0.5464, 0.12, 0.15, 0.26, 0.11, 0.29, 0.56, 4.1)^T$. The CPU time using a Pentium IV processor was $t = 5.5$ s and the distance between

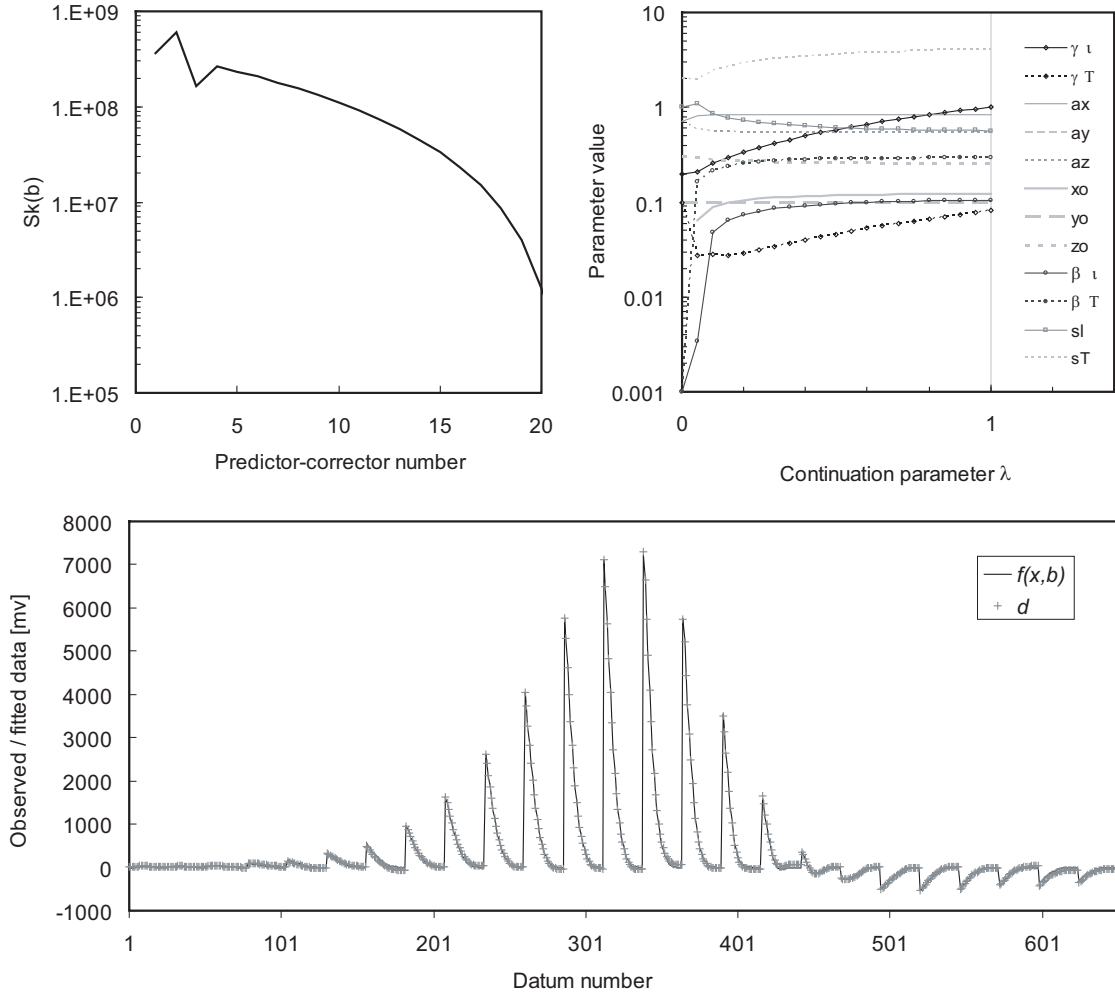


Fig. 21. Inversion results for a synthetic dataset corresponding to a stretched-exponential decaying magnetic spheroid. The dataset is contaminated with Gaussian noise having a SNR=25. Top left: the resulting cost function value as a function of the predictor-corrector iteration number. Top right: model vector evolution during the path tracking. After 20 predictor-corrector steps, the continuation parameter λ reaches the value 1, providing a final model to the initial non-linear problem. Bottom: 1-D plot of the data vector and final model response shows the quality of the fitting.

the starting and final models is $|\mathbf{b} - \mathbf{b}_0| = 2.35$ while the length of the continuation path Γ was 3.10. For comparison, the inversion of noise-free data matched the true model in $t = 4.9$ s with $|\mathbf{b} - \mathbf{b}_0| = 1.01$ (path length Γ was 2.72). The excellent fit between the synthetic data and the EMI spatiotemporal response of model \mathbf{b}_* is shown in Figure 21(bottom) as spatiotemporal series. The inversion result described here constitutes a reasonable gauge of the capability of the continuation algorithm to recover physically interpretable values for the stretched coefficient and decay rates.

a. Sensitivity to initial model

A robust non-linear inversion algorithm should provide a reasonable and consistent minimization of the cost function without relying on a good guess of the initial model. The sensitivity of the final model vector \mathbf{b}_0 to different starting models was assessed by changing the initial values of several model parameters and running the inversion using the noisy synthetic dataset described in the previous section.

On each trial, polarizabilities $\gamma_{L,T}$, stretch coefficients $\beta_{L,T}$ and decay rates $s_{L,T}$ were changed along with some elements of the spheroid axis direction vector $\hat{\mathbf{a}}$, and location \mathbf{r}_0 . The stopping conditions and predictor stepsize in section A were used for all the trials. For sake of simplicity on the graphical representation, we are just presenting continuation path projections $(\gamma_{L,T} - \lambda, \beta_{L,T} - \lambda, s_{L,T} - \lambda)$ that are related to the parameters describing intrinsic target features.

Path projections $\gamma_{L,T} - \lambda$ for six different starting models are shown in Figure 22(top). Initial longitudinal and transversal polarizabilities were varied almost two and one orders of magnitude, respectively. During each trial, the method was able to track a different path that consistently converged to a final model parameter value which is similar to the actual value (1.14, 0.08, respectively). Figure 22(middle) corresponds to path projections $\beta_{L,T} - \lambda$ for three different starting models. In this

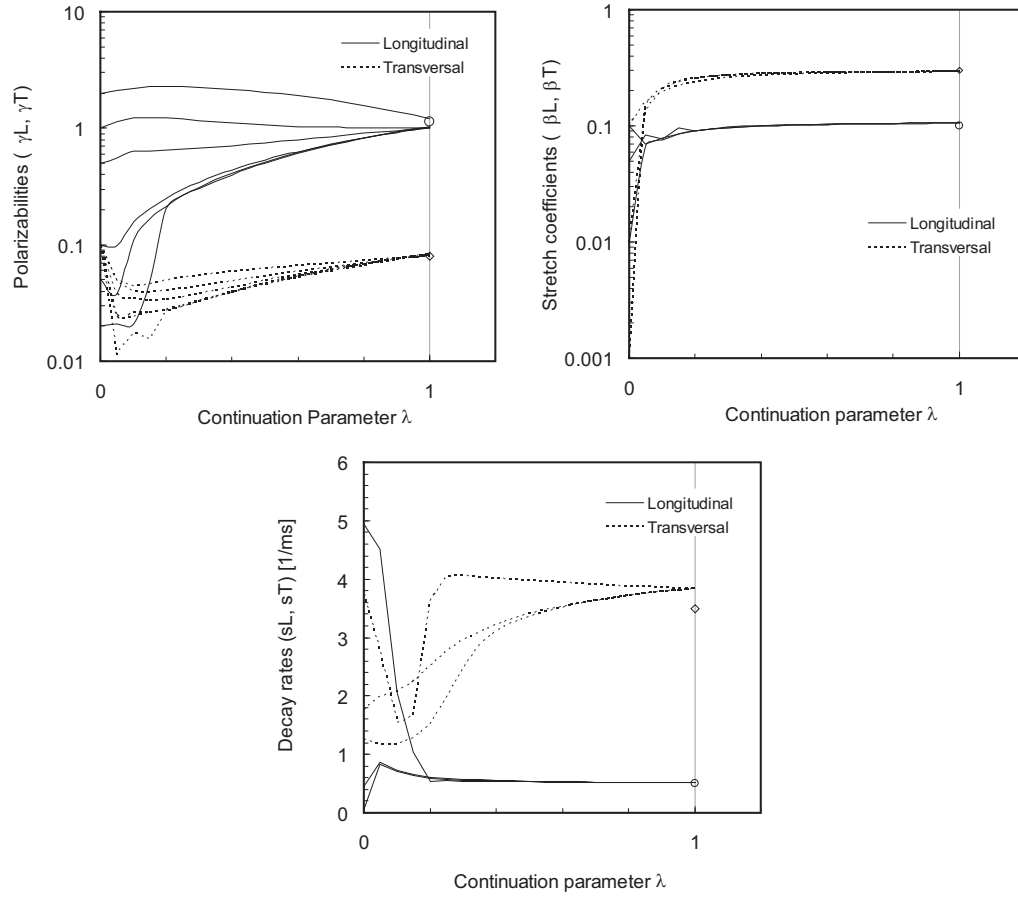


Fig. 22. Continuation path projections for different starting models using the synthetic dataset in Figure 21. In each plot, continuous line style corresponds to longitudinal model parameter, b_L while dashed line style corresponds to transversal model parameter, b_T . Open circles and diamonds at $\lambda=1.0$ represents the longitudinal and transversal model parameters used to generate the synthetic dataset, respectively. Continuation paths projected in the $\gamma_{L,T} - \lambda$ (top), $\beta_{L,T} - \lambda$ (middle), and decay rates $s_{L,T} - \lambda$ (bottom) planes show that the inversion algorithm consistently approximates the correct final model parameter value despite the initial model parameter is varied on one or two orders of magnitude.

case, the continuation path projection rapidly converges toward its final values in close agreement with the current values (open circle and open diamond at 0.1 and 0.3, respectively). It is evident that most of the continuation path along the stretch coefficients direction is constant and it is not sensitive to the initial values.

Finally, Figure 22(bottom) shows path projections $s_{L,T} - \lambda$ for three different starting models. The projected paths for the longitudinal decay rate exhibited similar convergence properties as in the case of the stretch coefficients. Also, the method was able to find a consistent final transversal decay rate value even though the continuation path projections depended on the initial value. The final value for the transversal decay rate is 9% off from the current value, possibly a consequence of the noise floor affecting the fastest decay component in the signal at late time. In general, the method was able to provide a final model vector in a consistent way even though the continuation path depends, in general, on the initial model.

b. Convergence and comparison with LM algorithm

Final cost function values and computing time for inversion of the synthetic dataset were compared using the new continuation method and a standard LM algorithm with cost function and model vector variation thresholds of 10^{-2} and 10^{-3} , respectively. In a later test, the cost function variation threshold was reduced to 10^{-3} . In both cases continuation parameter variation was $\Delta\lambda = 0.1$. The continuation algorithm improved the final cost compared to the LM algorithm.

B. 57mm M22 antitank ordnance

An inversion of actual field-measured EMI spatiotemporal data corresponding to a buried 57mm M22 ordnance is shown next. The 57mm M22 is an antitank artillery round ordnance 17.1 cm long and 5.7 cm in diameter. The EM63 multi-Rx data were

acquired with the ordnance buried at 0.2 m (depth below surface to the top of the ordnance). The axis of symmetry was inclined about 45° with respect to the vertical with the nose pointing down, striking in the positive x -direction.

The initial model vector \mathbf{b}_0 is the same as in the synthetic inversion but we used a longer predictor step size, $\Delta\lambda=0.1$, and the maximum number of GN corrector-step iterations was increased to ten. In this example, the depth-to-target parameter, z_0 was left adjustable. The continuation inversion was performed in 6 s of CPU time with $|\mathbf{b}-\mathbf{b}_0|=1.85$ and a continuation path length Γ of 5.42, indicating a considerable amount of curvature in the augmented model space (\mathbf{b}, λ) . The final model at $\lambda=1.0$ was $\mathbf{b}_*=(0.2506, 0.0012, 0.9796, 0.0, -0.2007, 0.1296, 0.0, 0.2928, 0.0137, 0.2458, 0.0732, 0.7946)^T$. There is a much stronger longitudinal eddy current response that lasts longer than the decay directed along the primary field. The inverted depth value matches the burial depth to the central portion of the ordnance.

The cost function (Figure 23, top left) shows again a very fast convergence within only 10 prediction steps. The parameter evolution (Figure 23, top right) during the path tracking shows considerable topological complexity in the tracked $(p+1)$ -dimensional path compared to the relatively simple path obtained for the synthetic inversion. Note that the polarizabilities and decay rates become less sensitive to changes in the continuation parameter λ during the last 3-4 predictor steps. This means that near the endpoint $\lambda=1$, there is a wider range of values that can be chosen to fit the data without affecting the cost function considerably. In other words, variations in those parameters, or combinations of them, produce similar EMI responses, i.e. there is degeneracy in the model parameterization. A certain amount of model equivalence, as a degenerate parameterization is otherwise known, is an inevitable characteristic of any non-linear geophysical inversion. The resulting final model response and observed data is shown in Figure 23(bottom). Notice the variable

quality of the fitting along the profile produced in this case by an underestimation of the inclination angle of the spheroid. However, amplitudes and decays are well reproduced for Rx locations inside the Tx loop (from datum 234 to 442).

C. 40mm MKII projectile

This ordnance is an artillery round 18 cm long and 3.3 cm in diameter. EM63 multi-Rx data were acquired with the top of ordnance buried at 0.14 m. The axis of symmetry was horizontal with respect to the ground surface with strike being negative x-direction (see Figure 19). In this case, we used $\Delta\lambda=0.05$ while the maximum amount of corrector iterations was 10. The continuation algorithm reduced the cost function value more than two orders of magnitude (Figure 24, top left). In this example, the inversion was performed in 90.45 s of CPU time with $|\mathbf{b}-\mathbf{b}_0|=1.518$ (path length Γ of 3.143). The continuation path showed that the transversal polarizability, the orientation unit vector component ax and the y -coordinate of the spheroid source caused most of the curvature during the path tracking (Figure 24, top right). The final model at $\lambda=1.0$ was $\mathbf{b}_*=(3.9\times 10^{-3}, 1.4\times 10^{-3}, -0.4228, -0.1659, 0.8909, 0.0146, -0.3111, 0.0875, 0.0652, 0.2182, 0.3577, 2.2314)^T$ with a high-quality data fitting in most of the Rx stations (Figure 24, bottom).

D. 60mm M49A3 mortar

This final example corresponds to a mortar (22 cm long and 6 cm diameter) buried at 0.2 m and oriented horizontally, as in the previous example. The shape of the response along the profile (see Figure 25, bottom) is quite similar to the one corresponding to the synthetic dataset (Figure 21, bottom). We used the same starting model and settings as in previous examples. The inversion reduced the cost function value by two orders of magnitude. Even though the reduction in cost function

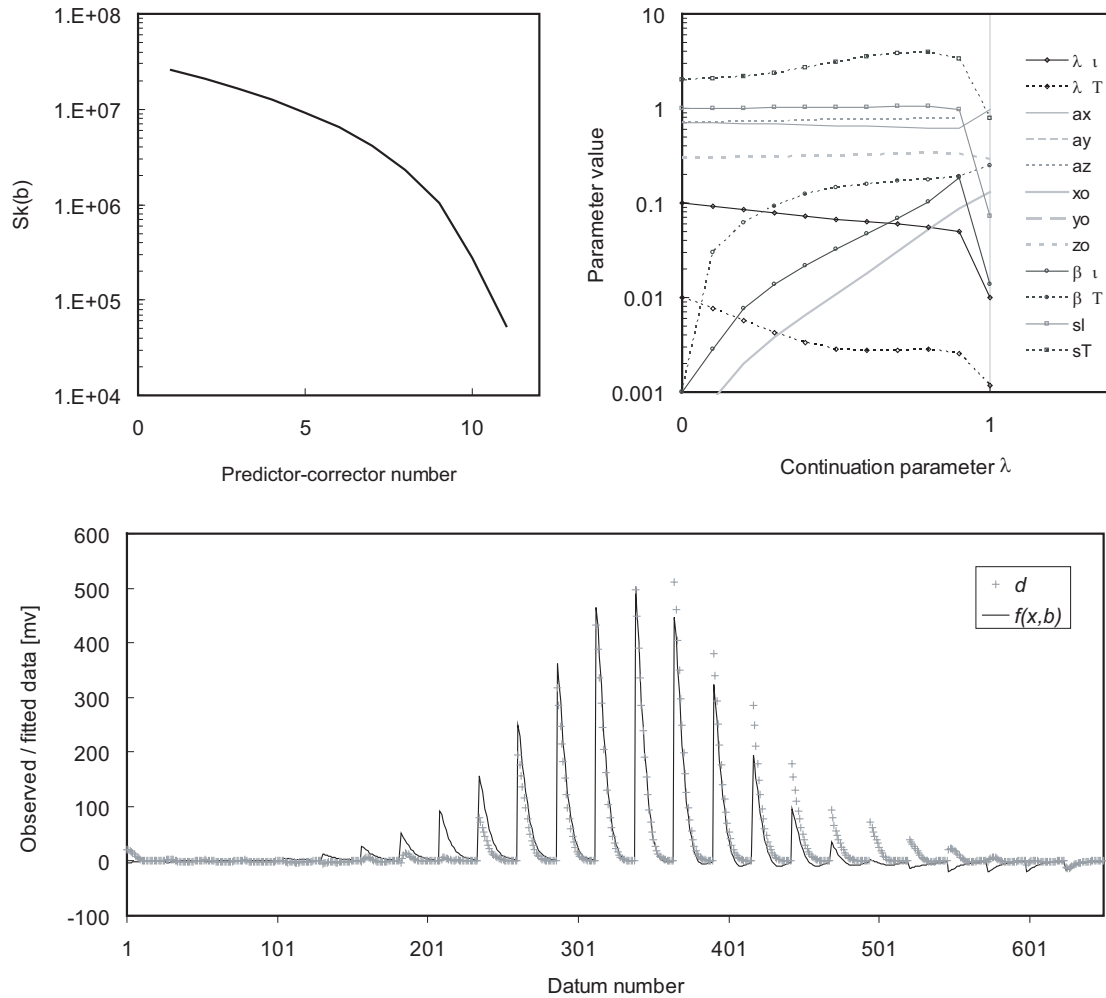


Fig. 23. Inversion results for a 57mm M22 antitank projectile. The ordnance is buried at 0.2 m (depth to the topmost part), inclined 45 degrees with respect to the horizontal (nose pointing down toward the positive profile direction). Top left: resulting cost function value as a function of the predictor-corrector iteration number shows almost 3 orders of magnitude in misfit reduction with almost exponential convergence near $\lambda=1.0$. Top right: model vector evolution during the path tracking. Bottom: data vector and the final model solution show an acceptable fitting.

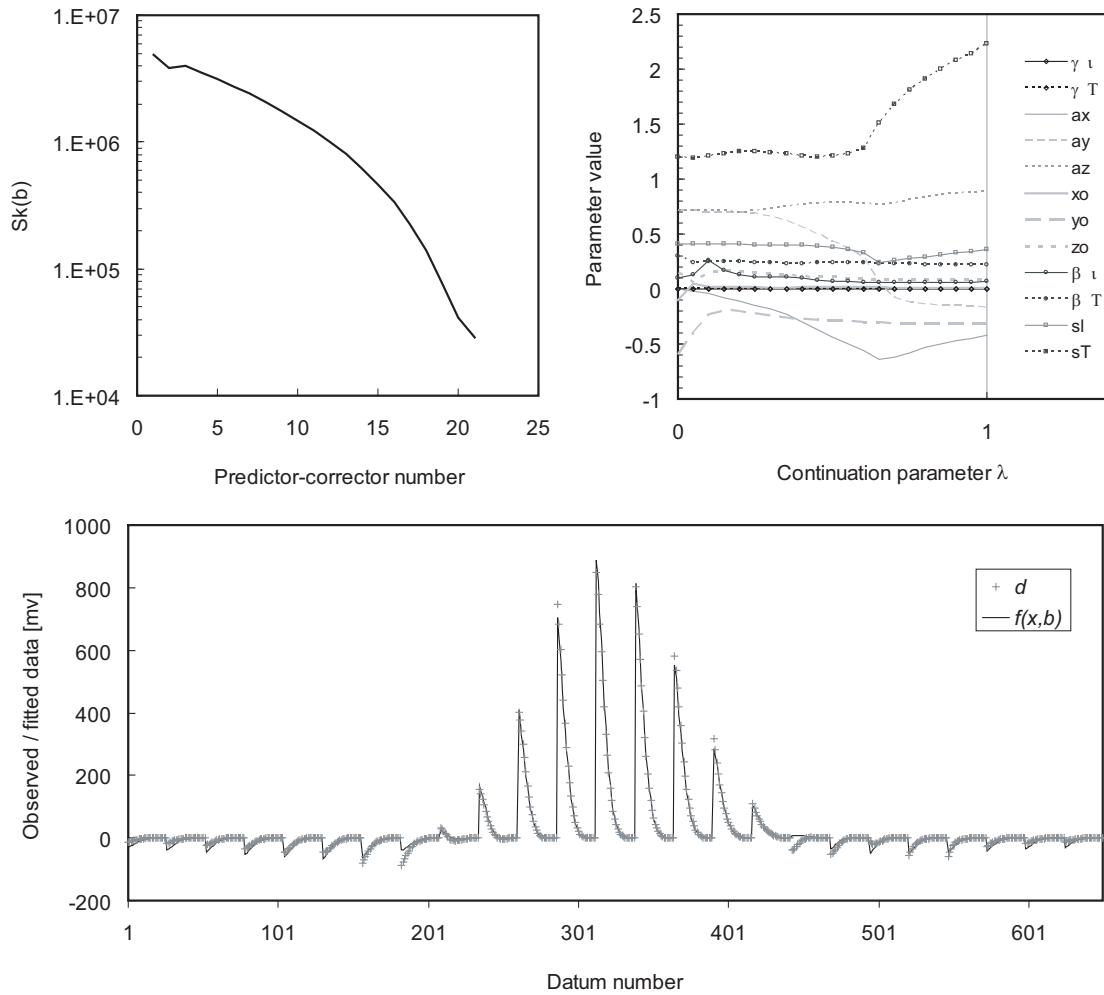


Fig. 24. Inversion results for a 40mm MKII projectile. The ordnance is buried parallel to the ground surface at 0.2 m (nose pointing to the start of the positive profile direction). Top left: cost function shows two orders of magnitude in misfit reduction. Top right: model vector evolution during the path tracking, showing a smoother variation of the parameters as a function of λ . Bottom: data vector and the final model solution. Notice that the quality of the fitting is better than in the case of the 57mm M22 ordnance; it reproduces all the features of the spatiotemporal response.

is not as smooth as in the other examples (see abrupt misfit change at predictor-corrector steps 10 and 11 in Figure 25(top left), convergence was achieved in 7.52 s of CPU time ($|\mathbf{b} - \mathbf{b}_0| = 0.562$, path length $\Gamma = 0.97$). The final model obtained at $\lambda = 1.0$ was $\mathbf{b}_* = (8.1 \times 10^{-3}, 6 \times 10^{-4}, 0.2815, 0.4940, 0.8227, 0.1775, -0.3491, 0.1970, 0.0509, 0.2937, 0.4938, 1.0665)^T$, providing a very good fit for the whole spatiotemporal data series. Figure 25(bottom) shows that both initial amplitudes and decays are very well represented by the final spheroid model.

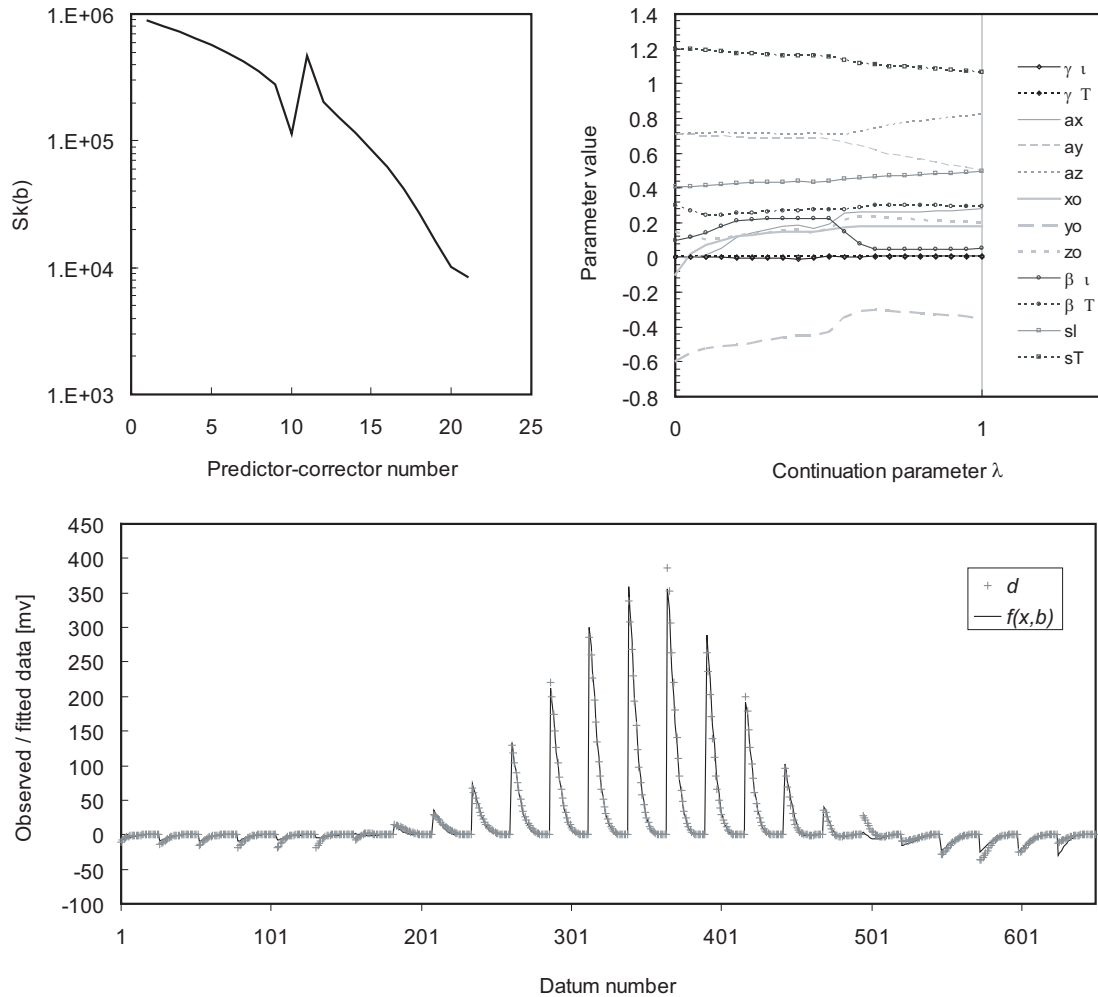


Fig. 25. Inversion results for a 60mm M49A3 mortar. The ordnance is horizontally buried at 0.2 m (nose pointing to the start of the positive profile direction). Top left: the cost function plot shows an abrupt oscillation at the 10th predictor-corrector step ($\lambda=0.5$). After three continuation steps, the misfit reduction trend stabilizes again. A reduction of two orders of magnitude in misfit is achieved. Top right: model vector evolution during the path tracking shows a direction change in most of the model parameters at $\lambda=0.5$. This can be interpreted as an increase in path curvature, reducing the effectiveness of the corrector step especially when the step-length along the continuation parameter is fixed. Bottom: fitting of the final model are comparable to the one obtained for the 40mm MKII projectile in Figure 24.

CHAPTER IV

UNSUPERVISED CLASSIFICATION OF CSEM DATA USING SELF-ORGANIZING MAPS

1. Introduction

Artificial neural networks (ANN) are a set of methods for signal or information processing based on the appropriate arrangement of simple processing units often called neurons. The core of each neuron consists of one or more mathematical operators performing linear or non-linear transformations of the input signals. Neurons in a network are connected in a deterministic way for data exchange, feedback, cooperation or competition to solve specific computational tasks.

The early developments on the concept of ANN were inspired by the description of brain cells, which started with the pioneering work of Golgi and Ramón y Cajal (1911). Brain cells, or neurons, consist of a main body, or soma, with two different kinds of ramifications; dendrites and axons. Dendrites exhibit an irregular surface with numerous branches while axons extend further away from the soma, exhibit a smooth surface and have few terminal ramifications. Signal reception occurs through the dendrite terminals while signal transmission is sent through the axon's terminals, in a process called synapses. Neural communication occurs between synapses and dendrites of connected neurons by complex electrical and chemical mechanisms. The number of synaptic receptors and synaptic terminals in a neuron are on the order of $10^3 - 10^4$. A human brain has 10^{11} neurons and 10^{14} inter-neuron connections sending information with characteristic times in the order of $10^{-3}s$ (Kohonen, 2000; Haykin, 1994).

In contrast, the basic unit of an artificial neuron is highly variable in terms

of number of inputs/outputs, connections, and computational processes. Artificial neurons store information in a set of real or complex numbers called a weight vector, \mathbf{w} ; a very simplified version of a synapse. A mathematical feature of all artificial neurons is the “interaction” between inputs \mathbf{x} and weights through a dot-product $\mathbf{w} \cdot \mathbf{x}$, convolution $\mathbf{w} \otimes \mathbf{x}$, or some distance measure between input and weights. The output is usually determined by a non-linear function of input, weights and a threshold value $f(\mathbf{w} \cdot \mathbf{x} - \theta)$. Also, artificial neurons can have feedback from the output to the inputs, providing memory for the processing of sequential inputs. The number of inputs in an artificial neuron is highly variable; from just a few up to hundreds or thousands. ANN can be simulated by software or implemented using integrated electronic devices. In the latter case, field programmable gate arrays (FPGA) consisting of $10^5 - 10^7$ logical gates are available, enabling the creation of ANN with $10^2 - 10^3$ neural units running in numerous parallel paths at the same speed as the board’s clock frequency.

Initial publications on ANN aimed to explain experimental discoveries on neuron functions and how the brain responds to simple stimuli. Currently, there are two well defined research paths; one aiming to incorporate knowledge into ANN models to reproduce more complex brain functions. The other focuses on making use of the ANN as a paradigm to process data. In neither case, has ANN been able to match the capability of biological systems; its accuracy, noise tolerance, generalization when exposed to new stimuli. The “low” processing speed of biological systems is compensated by an efficient parallel/distributed processes and effective way to represent and store stimuli information.

Artificial neurons using a dot-product operation were proposed first in the literature. First, as a switch with fixed threshold, binary inputs, and identical weight values (McCulloch and Pitts, 1943). Later, with variable weights wider output value range and learning rule (Rosenblatt, 1958), constituting the basic unit for single layer

perceptrons and modern ANN. The use of convolution (Caianello, 1961) tried to model the fact that biological neurons respond as a threshold unit by accumulating incoming stimuli over certain period of time.

A. Feed forward neural networks

Single artificial neurons can perform simple tasks such as some Boolean operations or classify data clusters which are linear separable (decision boundary is a line in a two-dimensional case or a hyperplane when dimension is larger than three) but fails to compute the Boolean exclusive-OR. By fully connecting the inputs of several neurons (neuron layer) to the input data and connecting their outputs to the inputs of one or more output neurons, the exclusive-OR operator and the classification of input patterns separated by complex decision boundaries was solved (Rosenblatt, 1958). The use of several neuron layers is referred in the literature as multilayer perceptrons (MLP) or multilayer feed forward neural networks (FFANN). FFANN are capable of approximating any measurable function to any desired degree of accuracy providing sufficient number of hidden neurons; this property is often referred in the literature as universal approximators (Hornik et al., 1989).

Training a FFANN can be treated as an optimization problem in which the neuron weights are adapted to minimize the difference between the network output and the expected output, provided some cost function. In other words, it is assumed that we have *a-priori* knowledge about the input data on a case-by-case or type-of-class basis and we *teach* the network to process input data patterns in an specific way. This approach is called *supervised learning*. A supervised ANN does not learn to store or *quantize* data, it learns to process it. In this case, to process it means also to store it temporary.

The use of FFANN in geology and geophysics is not new even though its appli-

cation has occurred slowly when compared with other applied disciplines. Reviews of the applications of ANN in these fields can be found in the literature (Poulton, 2001; Poulton, 2002; van der Baan and Jutten, 2000; Calderón-Macías et al., 2000). General tutorials (Guyon, 1991, for example) as well as a review ANN in electromagnetics (Christodoulou and Georgiopoulos, 2001) are also available. A detailed, but probably not complete, review of publications in different geophysical fields is as follows. In seismic enhancing fluid migration paths assessment (Ligtenberg, 2003), porosity thickness estimation (Trappe and Hellmich, 2000), fault analysis (Tingdahl and de Rooij, 2005; Boadu, 1998), cross-hole tomography velocity inversion (Nath et al., 1999), explosions location (Fedorenko et al., 1998), 1D or 2D velocity models inversion (Calderón-Macías et al., 2000; Zhang et al., 2002a), reflection data processing (Ashida, 1996; Dowd and Pardo-Iguzquiza, 2005; Strecker and Uden, 2002), source wavelet estimation (Wang and Mendel, 1992), reservoir characterization (An et al., 2001; Mercado Herrera et al., 2006), refraction interpretation (McCormack et al., 1993), and facies classification (Saggaf et al., 2003). In electromagnetics, magnetotelluric time-series analysis (Manoj and Nagarajan, 2003), mineral prospectivity mapping (Brown et al., 2000), 3D magnetotelluric inversion (Spichak and Popova, 2000; Zhang and Paulson, 1997), 1D multi-frequency airborne electromagnetic data inversion (Ahl, 2003), resistivity soundings inversion (Calderón-Macías et al., 2000; El-Qady and Ushijima, 2001), CSAMT geoelectrical parameters of fault zones (Spichak et al., 2002), buried utilities detection with GPR (Al-Nuaimy et al., 2000). In potential fields, sub-basalt gravity imaging (Fitzgerald and Bean, 2001), magnetic data analysis (Bescoby et al., 2006), Euler deconvolution (Mikhailov et al., 2003). In near-surface applications, groundwater transmissivity (Mukhopadhyay, 1999), prediction of hydraulic conductance, (Coppola et al., 2005a), prediction aquifer water level (Coppola et al., 2005b), ground-water remediation analysis (Johnson and Rogers, 1995), soil water retention

prediction (Koekkoek and Boeltink, 1999), and soil hydraulic parameters estimation (Schmitz et al., 2005). In well logging, neutron data interpretation (Aristodemou et al., 2005), fracture frequency prediction (Fitzgerald et al., 1999), rock properties (Bhatt and Helle, 2002), thermal conductivity prediction (Goutorbe et al., 2006), porosity and permeability prediction (Helle et al., 2001), borehole rocks lithology estimation (Benaouda et al., 1999), resistivity log interpretation (Zhang et al., 2002b; Zhang and Zhou, 2002), AVO inversion (Russell et al., 2002), and reservoir modeling (Basu et al., 2004).

B. Self organizing maps

Unsupervised or self-organized learning consists of training an artificial neural network without input about the function to be learned, clusters the input data belongs to, or the minimization of some predetermined cost function. The goal is to learn significant patterns from the input data and mapping them at specific locations on the network (Haykin, 1994).

Self-organizing maps (SOM) are another type of ANN in which neurons are arranged in a regular lattice, competing and associating themselves each time a new input pattern is presented during training. SOM were inspired by the organization and functions of the cerebral cortex. Cerebral cortex is the uppermost part of the brain and it contains important sensory functions grouped in a tiered fashion. Visual and somatosensory signals trigger neuron responses at specific cortex areas (Kohonen, 2000). More striking is the fact that some of those areas are spatially related in the cortex as the sensory organs are located in the human body, providing some sort of anatomical projection map. Cortical maps are associated to specific features, anatomical sensory excitations or abstract features. Current experimentation on functional NMR or magnetoencephalography shows that advance abstract associ-

ations and categorizations takes place in the cortex (Low et al., 2003, for example of the latter).

SOM provides a way of representing multidimensional data in lower dimensional spaces; usually one or two dimensions. The process of reducing the dimensionality of the inputs, is essentially a data compression technique called *vector quantization*. SOM creates a network that stores information in such a way that any topological relationships within the training set are maintained. Because the input dataset is the target of the SOM during the training process, there is no need to introduce output targets; the inputs are the output targets themselves. Training a SOM is like a self-teaching process or *unsupervised learning*. Unsupervised learning is useful to discover underlying structure of the data; data encoding, compression and transformation, data clustering. The latter, is the objective of the present chapter. In contrast to a FFANN, an unsupervised ANN does not learn to process data but to stored it in a compact representation by clustering.

Literature on SOM is exceptionally large in research fields such as biophysics, data/web mining, robotics, just to mention a few. In contrast, the amount of SOM application in geosciences is relatively small, almost non-existent when compared with the amount of publications on FFANN. Most of the SOM papers on geophysics deals with seismic applications: unsupervised seismic facies analysis from a deep water oilfield (Castro de Matos et al., 2007), classification of seismic salt-contaminated velocities (Lin et al., 2002), fracture spacing and elastic parameter prediction (Klose, 2006), Facies extraction using post-stack seismic attributes (Strecker and Uden, 2002) or from core description (Hsien-Cheng et al., 2002), and classification of multiple reflections (Essenreiter et al., 2001). Non-seismic applications include geo-referenced data processing (Bação et al., 2005) and determination of depth and σt product from multi-frequency profiles (Poulton et al., 1992a; Poulton et al., 1992b). A few books

specialized on SOM can be found (for example; Kohonen, 2001) while several good book chapters provides introductory reviews about the SOM algorithms (Haykin, 1994; Christodoulou and Georgiopoulos, 2001).

C. Classification of time-domain CSEM responses

As we saw in chapter III, section C, time-domain CSEM responses are particularly “corrupted” by the response of the background (or host, which in the case of very near surface, regolith or saprolith) at early times. The amount of corruption mainly depends on the electrical conductivity of the host and the size of the target. For small volume targets, background dominates early time response so it is not possible to visually assess the presence of the target by observing the secondary field decay or spatial profile. For targets with considerable volume, the background response is a small part of the response amplitude. Also, experiments performed using a Geonics EM63, and numerical simulations (Stalnaker et al., 2006) shows that at late time (i.e. when the EM diffusion length is comparable to the target size), some degree of target-host interaction, possibly due to galvanic currents in the host medium that dissipates charge accumulations on the target-host interface, also contaminates target responses.

The approach of introducing a simple physics-based forward model to estimate target parameters, like the magnetic spheroid model described in chapter III, section B, is effective as long as we have removed the background response prior to the inversion. In our experiments background response was determined right before burying the target. However, in real case scenarios, background responses should be determined by averaging responses at locations close to the potential target that are considered target/clutter-free or by predicting it by using previous measurements along the profiling direction. This detail is often ignored in the literature on cur-

rent UXO research. Target responses are acquired in *free space* conditions by using non-metallic platforms several feet above the ground or by digging an open pit and placing the target on it without burying it.

The challenge of performing an effective background-clutter-target discrimination in normal survey conditions is studied in this chapter. The idea is to process and learn from the simple CSEM transients and spatiotemporal responses that contains pure background, background and clutter, background and targets to see if it is possible to produce data clusters that can be used for classification purposes.

D. Chapter organization

This chapter is organized by first reviewing the self-organizing maps; their characteristics, algorithms, and output visualizations followed by an up-to-date description of self-organizing maps for time series; Temporal Kohonen maps and the theory of recurrent SOM. From the latter, a whole section will be devoted to merge SOM which constitutes the algorithm used to process Geonics EM63 spatiotemporal dataset. Central loop transients and spatiotemporal data is described following the pre-processing steps performed over the data. Results and discussion of synthetic and real data are presented to show the SOM and MSOM capabilities and finally, the results from Geonics EM63 central loop and spatiotemporal datasets in which the geological background response is included are presented and described.

2. Methods

A description of the basic Kohonen SOM algorithm follows. Neurons close in the SOM space are close in the input space. This learning behavior leads to partition the input space into convex regions. Weight vectors belonging to the same convex

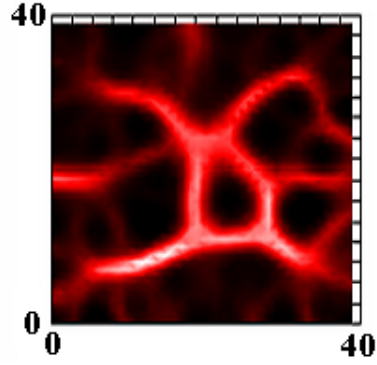


Fig. 26. Visualization of convex regions in a SOM. In this example, a 40×40 neuron SOM was trained using 100 input vectors representing colors in the RGB system. Each convex region corresponds to a specific color/hue.

region are closer to its mean weight vector than to the mean weight vector of any other convex region. The formation these convex domains is similar to the definition of what is called *Voronoi tessellation*. Since a SOM maps the input vectors into the network space, the formation of convex regions represents clusters of input data, like in the example of clustering color RGB indexes (Figure 26). Training a SOM can be viewed as minimizing a cost function that leads to a weight space ordering by creating partitions. Because a cost function is, in general, discontinuous between regions, it can not be defined as an energy function (Erwin et al., 1992).

A. Kohonen SOM

in the case of a self-organizing map Q with N neurons arranged on grid of nodes with indexes i, j the cost function can defined as the sum of the weighted norm of the difference between weight (\mathbf{w}) and input vectors ($X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M)$). Using Euclidean norm, the cost function $S(Q, \mathbf{x})$ is:

$$S(Q, \mathbf{X}) = \frac{1}{N} \sum_{i \in Q} \sum_{j=1}^M h_{ij} \|\mathbf{x}_j - \mathbf{w}_i\|^2. \quad (4.1)$$

The neighborhood function h_{ij} determines the contribution of each vector difference to the total cost and it is measured with respect to the best matching unit (bmu) \mathbf{w}_I neuron of the input \mathbf{x}_j . The usual choice for the neighborhood function is an unnormalized Gaussian:

$$h_{i,I}(n) = \exp \left[-\frac{\|\mathbf{r}_i - \mathbf{r}_I\|^2}{\sigma(n)^2} \right]. \quad (4.2)$$

Where \mathbf{r}_j corresponds to the distance between any network unit \mathbf{r}_i and the bmu, \mathbf{r}_I . Notice that the value of the neighborhood function depends on the variance $\sigma(n)$ which in turns decreases with the iteration (or epoch) number. As the learning process goes on, the neighborhood shrinks. During the final epochs, the neighborhood function converges to a Dirac delta. Another neighborhood function described in the literature is the unnormalized Mexican hat:

$$h_{i,b}(n) = \left[1 - \frac{\|\mathbf{r}_i - \mathbf{r}_I\|^2}{2\sigma(n)^2} \right] \exp \left[-\frac{\|\mathbf{r}_i - \mathbf{r}_I\|^2}{2\sigma(n)^2} \right]. \quad (4.3)$$

Which can also be described as the difference between two Gaussian functions with different variances. One of the SOM learning methods is called stochastic gradient. It consist of two steps each time an input pattern is presented to the network: i) competition to determine the *bmu*, and ii) association to spread information from the current input vector across the network. In the first step the *bmu*, \mathbf{I} , with network coordinates (I_i, I_j) will be the neuron whose weight vector is the closest to the input vector,

$$\mathbf{I} = \operatorname{argmin}_{(i,j \in Q)} \|\mathbf{x}(m) - \mathbf{w}(n)\|. \quad (4.4)$$

Position of the *bmu* determines the neighborhood center and so the neurons that will learn about the input vector. In the second step, weight update is a function of the

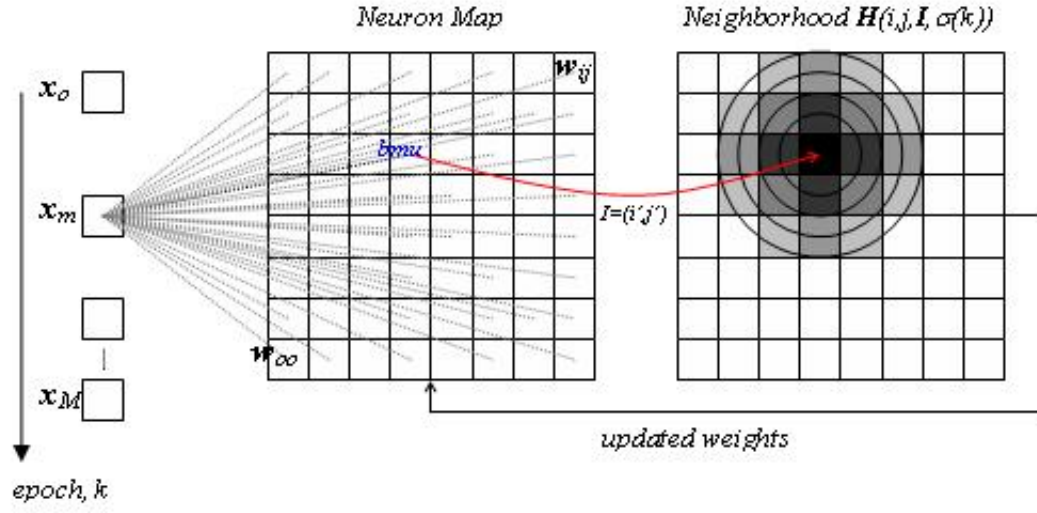


Fig. 27. Sketch of a self-organizing map learning process. The neuron network consist of a 2-dimensional array of neurons characterized by local coordinates (i, j) and a weight w_{ij} . During the k -th epoch, an input vector \mathbf{x}_m is presented to the network and the neuron with the closest distance to the input vector is taken as the *bmu*. Neighborhood function map $H(i, j, I, \sigma(k))$ at the left hand side, represented by different grayscale intensities, determines the strength of association between neurons.

distance to the current *bmu*, expressed through the neighborhood function $h_{i,I}(n)$

$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + \gamma(n)h_{i,I}(n)(\mathbf{x}(n) - \mathbf{w}_i(n)), \quad (4.5)$$

where $\gamma(n)$ is a scalar called learning rate bounded on the interval $(0, 1]$. The *bmu* keeps the larger weight correction ($h_{I,I}(n) = 1$), getting tuned in the direction of the input vector. The association effect takes place at the neighboring nodes but to a lesser degree. Exposing the network to similar input patterns will produce further activations in the same neighborhood, thus clustering the data.

The stochastic gradient method initializes the weight vectors to small random numbers. Input vectors \mathbf{X} are presented to the network in a random fashion to provide learning from all the possible clusters without bias for an specific class. Randomiza-

tion can be achieved by picking up the input vector at random or sequentially, using a randomly sorted database. Using a sequential scheme to present the inputs, one iteration over the dataset \mathbf{X} is referred to as an *epoch*. For the network to learn about the input patterns, a certain number of epochs is required. Cycling several times over the input dataset (also known as *incremental training*) is very common in ANN training methods. Figure 27 visualizes a SOM as a grid in which each square represents a neuron. Each input pattern should interact with each neuron in order to determine the *bmu*. A neighborhood function can be calculated each time a new pattern is presented or after each epoch. The latter case saves computational time and it does not seem to produce negative effects on the convergence rate of training. If for each epoch, a neighborhood map twice the size of the SOM with the *bmu* located at its center is computed, weight update will not require additional calculations but a simple subset extraction with the center of the neighborhood map located at the current *bmu* position.

Every time an input vector is presented, the algorithm calculates $S(Q, x)$, equation (4.1) and keeps track on the position of the *bmu*. The idea is to monitor the convergence of the training, to implement training stopping conditions, and to verify that each neuron have succeeded in becoming a *bmu*, assuring a good data representation. This is important because there is no analytical way to calculate the optimum values and time dependency of $\sigma(n)$, η (equations 4.2 and 4.5, respectively), and the SOM size. For a given dataset, two or more training sessions are needed to optimize the learning process. In the present algorithm, the average of the cost function per epoch (n) per neuron $R(S_n, \mathbf{X})$ is the quantity monitored

$$R(S_n, \mathbf{X}) = \frac{1}{M} \sum_{j=1}^M S_n(Q, \mathbf{x}). \quad (4.6)$$

The training is stopped when $R(S_n, \mathbf{X})$ is smaller than a certain fraction of its initial value $R(S_n, \mathbf{X}) < \delta R(S_1, \mathbf{X})$ ($\delta \approx 10^{-4} - 10^{-6}$) and it does not change over several number of epochs, usually 5 – 10.

a. SOM visualizations

After a SOM has been trained, the network has stored the training dataset on M neurons. If a new input vector is presented, the SOM output is very simple; a *bm*. However it is important to be able to visualize all the possible SOM outputs in terms of the knowledge available from the original training datasets or in terms of the way weights were organized. Original work on SOM (Kohonen, 2000) showed the trained map with each unit associated with a specific quality, class, etc. Recent advances in SOM have provided valuable methods for visualization of data clusters. Examples are the U and U^* matrices (Ultsch, 2003b; Ultsch, 2003a), a U^*F matrix (Moutarde and Ultsch, 2005), and vector fields (Polzlbauer et al., 2005). In this section, a primer on U -matrix, and U^* -matrix will be presented.

For a given neuron i with weight vector \mathbf{w}_i , a measure of similarity with respect to its immediate neighbor neurons $D(i)$ up to certain network distance can be written as

$$U(i) = \sum_{j \in D(i)} d(\mathbf{w}_i - \mathbf{w}_j), \quad (4.7)$$

where $d(\mathbf{a}, \mathbf{b})$ is a distance based on the same metric used for training the network. In the present work, we modify equation (4.7) by taking the median of the distances to the neighbors. In this way, U -matrix values are kept small, preserves large variations in U -values while removing small fluctuations. A plot of $U(i)$ in the local network coordinates as an intensity or three dimensional map is called *U-matrix* (Ultsch, 2003b). In a square network lattice, the one used in this work, there are 4 immediate

neighbors. There is a trade off between the amount of details the U-matrix can provide and the size of the neighborhood; usually the first 4 or 8 neighbors are sufficient to obtain a good map.

Notice that U-matrix provides a picture of the local distance structure of the patterns stored on the network, and so displays the local distance structure of the training dataset. The structure is highlighted by the presence of valleys and ridges, as depicted in Figure 26, section 1 of this chapter. Valleys, associated with low U values, contain neurons whose weights are close together, high U values contain neurons with weights that are distant from the weights of its neighbors. Proper labeling of the *bmj* on top of the U-matrix shows that the projections of the input vectors cluster in the bottom of the valleys, meaning that the ridges are the U-matrix representation of the cluster boundaries. An additional property of the U-matrix is that outliers in the input dataset become isolated in small depressions (or funnels). The idea is to provide classification of input vectors since the U-matrix interpretation helps to avoid misclassification produced by an incomplete training dataset or “contaminated” patterns. This is especially relevant when dealing with high-dimensional input spaces in which it is impossible to assess these special cases.

One drawback of the U-matrix is that local distances in dense regions of the weight space are often related to the same cluster in the data space, meaning that they will not provide information about cluster boundaries. The opposite effect occurs in low-density regions of the weight space where a cluster boundary is more likely to be present. Scaling the U-matrix based on an estimate of the input data density about each neuron weight helps to improve the visualization of cluster boundaries. Density estimation may be estimated by dividing the number of vectors located at distances within the radius of the hypersphere, centered at the weight vector, by the volume of the hypersphere. If the same hypersphere radius is used, it is sufficient to count the

number of vectors to get some proportional measure of density. The selection of the radius is critical to avoid under or overestimation of data density, especially if the data clusters overlap. The hypersphere radius was calculated using the cumulative density function (CDF) of distances between input vectors ($\|\mathbf{x}_i - \mathbf{x}_j\| \in \mathbf{X}$). The distance corresponding to a cumulative probability of 20% (Pareto radius) is chosen as the reference radius (Pareto spheres). It has been shown that this method provides a good density estimator (Ultsch, 2003a).

Using the Pareto radius, a matrix of data density can be estimated by counting how many input vectors $P(\mathbf{w}_i)$ are within a Pareto radius of each weight vector \mathbf{w}_i . $P(\mathbf{w}_i)$ plotted at the network coordinates is called *P-matrix* (Ultsch, 2003b). The P-matrix is scaled using the following formula

$$P_{scaled}(i) = \frac{p(i) - \langle p \rangle}{\langle p \rangle - p^{max}} + 1, \quad (4.8)$$

where $\langle p \rangle$ is the average density estimation, and p^{max} the maximum value within the P-matrix. It is easy to see that at neuron locations with data vectors density below average produces scale factor below 1. For locations with high density of data vectors, the scale factor are greater than 1. The U*-matrix ($U^*(ij) = P_{scaled}(ij)U(ij)$) is an improved version of the U-matrix which incorporates local input data density estimates (Ultsch, 2003b); it provides flattened cluster regions, sharpens the cluster boundaries, and compresses the amplitude range of the U-matrix.

B. SOM for time series

Standard self-organization requires static input vectors to provide topographical representations. These vectors may contain raw/pre-processed data or features which represent important discriminatory information. Often times, input data consists of time-dependent, spatial profiles, maps, or spatiotemporal series with high dimension-

ality which are difficult to process and represent with SOM. The problem is often referred to as the *curse of dimensionality* in inverse theory. Several approaches to attack this issue are available in the literature, some of them will be described in the next two sections. The key point here is to process portions of the input data sequentially by defining an activation function that not only learns from the current “slice” of information but also keeps knowledge about the time/spatial characteristics of previous slices.

There are several approaches to process sequential data, starting with moving windows of certain fixed length (Kangas, 1990), storing previous activations by propagating them laterally on the neuron network (Wiemer, 2003), recurrent processing of the spatial/temporal slices by preserving previous activation features for the calculation of the current activation (Chappel and Taylor, 1993; Strickert and Hammer, 2005; Varsta et al., 1997). For a general review the reader can refer to (Barreto et al., 2003). In the case of recurrent SOM, neurons store certain amount of information about the activation produced on the previous input slice in a scalar or vector representation called *context*. Activation of a recurrent SOM will take into account both weights and contexts to incorporate present and previous data inputs. In the next section, a generalized recurrent SOM model will be presented following the work of Hammer (Hammer et al., 2004), followed by a more detailed description of the merge SOM used for precessing Geonics EM63 spatiotemporal data.

a. Generalized recurrent SOM

Given a finite set of training data $T = [\mathbf{X}_1, \dots, \mathbf{X}_p]$, where \mathbf{X}_p corresponds to an input data pattern, the idea is to obtain a neural map that represents the input data as accurately as possible (Hammer et al., 2004) and, at the same time, to store enough information variability so the network can accommodate new input

patterns. Following the nomenclature used for SOM in section A above, an input pattern $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M)$ is defined as a set of vectors of length l that constitutes a time or spatiotemporal slice. Also, there is a network of N neurons. For the purposes of the present work, let us assume that each slice \mathbf{x}_t is a vector corresponding to a spatial response at time t . The interactions of the input pattern at time t in neuron i can be expressed by a dynamical equation

$$\tilde{d}_i(t) = \alpha d(\mathbf{w}_i, \mathbf{x}_t) + (1 - \alpha) d_r(\mathbf{c}_i, \mathbf{C}_t), \quad (4.9)$$

that involves both weights \mathbf{w} and contexts \mathbf{c} . The scalar $\alpha \in (0, 1)$ is the context parameter. Equation (4.9) measures interaction at neuron i using some measure of similarity, $d(\mathbf{a}, \mathbf{b})$, between input slice and weights, and between current context and context descriptor \mathbf{C}_t . Context descriptor can be defined as a vector function,

$$\mathbf{C}_t = F(\tilde{d}_1(t-1), \dots, \tilde{d}_N(t-1)), \quad (4.10)$$

This generalized context descriptor $F : \mathfrak{R}^N \longrightarrow \mathfrak{R}^r$ is a function that extracts the relevant information from the map at the previous time step (Hammer et al., 2005). Several recurrent SOM found in the literature can be described as special cases of equations (4.9) and (4.10). In the case of $r = N$, $F = id$, \mathbf{c} an unit vector of the current activation, and d_r a dot-product operation, the recurrent SOM leads to what is called a Temporal Kohonen Map (TKM) (Chappel and Taylor, 1993) whose dynamical equation and learning rule can be written as

$$d_j(\mathbf{X}, \eta) = \sum_{i=1}^M \eta(1 - \eta)^{M-i} \|\mathbf{x}_i - \mathbf{w}_j\|^2. \quad (4.11)$$

$$\mathbf{w}_j(t) = \mathbf{w}_j(t-1) + \gamma h_{j,I}(\mathbf{x}_i - \mathbf{w}_j(t-1)), \quad (4.12)$$

Memory parameter η is bounded in the interval $\in (0, 1)$. The winner after presenting a pattern X is the neuron closer to the average of recent slices. Memory is provided by the decay factor $(1 - \eta)^{M-i}$ which models the contribution of each time slice to the sum; the earlier a slice is from the current time, the weaker is its contribution. Memory in KTM decays exponentially with the time. This type of memory effect is referred in the literature as *leaky integration*.

A variation of KTM is the Recurrent SOM (RSOM). The dynamical equation measures the norm of the leaky integration (Varsta et al., 1997);

$$d_j(\mathbf{x}, I_t) = \left\| \sum_{i=1}^t \eta(1 - \eta)^{i-1} (\mathbf{x}_i - \mathbf{w}_j) \right\|^2, \quad (4.13)$$

providing a broader context storage because directions, instead of distances are stored.

A Hebbian learning rule

$$\mathbf{w}_j(t) = \mathbf{w}_j(t-1) + \gamma h_{j,I} \sum_{i=1}^t \eta(1 - \eta)^{i-1} (\mathbf{x}_i - \mathbf{w}_j(t-1)), \quad (4.14)$$

updates the weights in the direction of the average for the previous time slices. In both KTM and RSOM, the amount of input pattern encoding stored on the weight space depends on the memory parameter η .

In the case when $r = N$, neurons contain both weight and context vectors which represents the network activation in the previous time slice. This type of network is called recursive SOM (RecSOM) (Voegtlin, 2002). The dynamical equation and context descriptor are given by

$$\tilde{d}_i(t) = \alpha_1 \|\mathbf{x}_t - \mathbf{w}_i\|^2 + \alpha_2 \|\mathbf{C}_t - \mathbf{c}_i\|^2, \quad (4.15)$$

$$\mathbf{C}_t(\mathbf{x}, \mathbf{x}) = (\exp(-d_1(t-1)), \dots, \exp(-d_N(t-1))). \quad (4.16)$$

Notice the the dimensionality of \mathbf{w} and \mathbf{c} are not necessarily the same. In this way,

RecSOM preserves information about the dynamics of the whole neuron network in previous time slices and measures its similarity with the current context \mathbf{c}_t . Learning is obtained by using a similar update rule for weights and context, as in the KTM (equation 4.12)

$$\begin{aligned}\mathbf{w}_j(t) &= \mathbf{w}_j(t-1) + \gamma h_{j,I}(\mathbf{x}_i - \mathbf{w}_j(t-1)) \\ \mathbf{c}_j(t) &= \mathbf{c}_j(t-1) + \gamma h_{j,I}(\mathbf{C}_j - \mathbf{c}_j(t-1)).\end{aligned}\tag{4.17}$$

One drawback of RecSOM is that the size of the context descriptor \mathbf{C} and context \mathbf{c} grows with the size of the network N , becoming a computational burden for large scale problems.

In the case of $r = n$, $d_r = d$, and context descriptor based on a linear combination of weights and context of the previous winner neuron defines what is called Merge SOM (MSOM) (Strickert and Hammer, 2005). Because MSOM was chosen to process Geonics EM63 spatiotemporal data, a more general description of its properties and algorithm will be provided in the next section.

C. Merge SOM

MSOM appeared very recently as a data processing tool that performs better than KTM and RSOM and is much less computationally expensive than RecSOM. MSOM have been used for DNA nucleotide splice site detection, speaker identification, binary automata and Reber Grammar detection (Hammer et al., 2005; Strickert and Hammer, 2005). It can be useful for time series inspection, clustering or *a-posteriori labeling* classification or regression since the network sorts the input patterns based on the winners of previous time slices, which incorporates temporal statistics (Hammer et al., 2005).

a. MSOM algorithm

Using the same notation used for recurrent SOM, the dynamical equation for MSOM is defined as a linear combination of Euclidean distances for weights and context vectors

$$\tilde{d}_i(t) = (1 - \alpha) \|\mathbf{x}_t - \mathbf{w}_i\|^2 + \alpha \|\mathbf{C}_t - \mathbf{c}_i\|^2, \quad (4.18)$$

For an input pattern \mathbf{X}_p , the activation of the network at time t after presenting the time slice \mathbf{x}_t is determined by the *bmu*

$$\mathbf{I}_t = \operatorname{argmin}_{(i,j \in Q)} \left\{ \tilde{d}(t) \right\}. \quad (4.19)$$

Context descriptor (also called *merge context*) is then calculated using a linear combination of weights and current context of the bmu

$$\mathbf{C}_t = (1 - \beta) \mathbf{w}^{I_{t-1}} + \beta \mathbf{c}^{I_{t-1}}, \quad (4.20)$$

where $\beta \in (0, 1)$ is the merge parameter. The initial condition $\mathbf{C}_{t=0} = \mathbf{0}$ is used each time a new pattern \mathbf{X} is presented to the network. Learning is achieved by using equations (4.17) for RecSOM with a Gaussian neighborhood function $h_{i,I}$ (equation 4.2). It has been shown (Strickert and Hammer, 2005) that in the final learning steps, when the neighborhood function becomes a Dirac delta and cooperation between neurons not longer takes effect, learning provided by equations (4.12) and (4.17) converges to optima weights and contexts vectors given by

$$\mathbf{w}_{opt(t)} = \mathbf{x}_t, \quad (4.21)$$

and

$$\mathbf{c}_{opt(t)} = \sum_{i=1}^{t-1} \beta (1 - \beta)^{i-1} \mathbf{x}^{t-i}, \quad (4.22)$$

respectively. Notice that the optimum weight $\mathbf{w}_{opt(t)}$ converges to the input time slice while the context $\mathbf{c}_{opt(t)}$ converges to the leaked integration of the input pattern with memory constant equal to the merge parameter β . This result means that a MSOM with sufficient learning time, is able to reproduce the input patterns in an exact way. Also, if the MSOM network is large enough, it is possible to obtain good representations of new input patterns.

The learning process of the MSOM algorithm starts by defining a network of N neurons arranged in a fixed, regular grid. Each neuron will store a tuple (\mathbf{w}, \mathbf{c}) of randomly initialized values. The value of the merge parameter β is usually set to 0.5. Selection of the context parameter α strongly influences the learning process and its optimal value is difficult to determine. It has been suggested that α should be adaptive, starting with a small value, $\exp(-10)$ and then updated after each input pattern is presented by measuring the total *entropy* of the network. The value of α is increased in steps of 0.05 until a clear trend in entropy variation is achieved. Afterwards, if entropy increases, α is increased by a factor of 1.2, if entropy decreases, α is reduced in 0.05 (Hammer, personal communication). Experience on processing data has shown that convergence takes place for weights and contexts in different stages during learning; first for the weights, then for the context representations (Strickert and Hammer, 2005). This effect has been confirmed in our research after numerous MSOM learning tests using synthetic and real datasets. Good learning results has been obtained by using positive, linear increment on α , starting on $0.01 - 0.001$, and ending in the range $0.1 - 0.3$. Updates occurred after each learning epoch. For each input pattern \mathbf{X} , the merge context is initialized and the sequence of time slices is presented to the network. Each time slice will activate a *bmu*, which updates the merge context \mathbf{C}_t . Learning occurs in the whole network by updating weights and contexts with strength determined by the neighborhood function. An input pattern

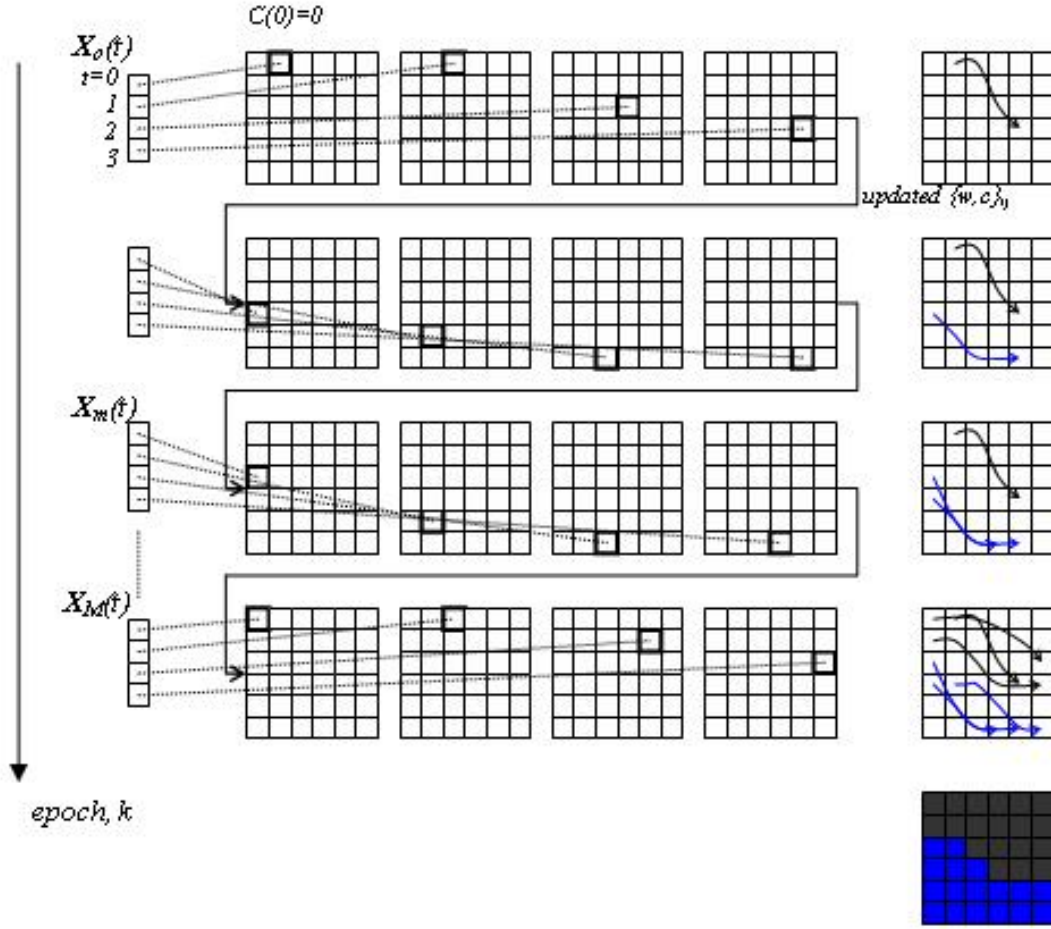


Fig. 28. Sketch of a merge self-organizing map learning process. During the k -th epoch, an input spatiotemporal pattern $\mathbf{X}_0(t)$ is presented in sequential steps to the network. For each input pattern, tracking of the updated tuple (\mathbf{w}, \mathbf{c}) and the merge context vector \mathbf{C} is performed, producing a path of *bmus*'s. The set of paths is then used for pattern clustering.

will contain a set of p *bmu*'s that defines its activation path. (see Figure 28). After the learning process has converged, the MSOM will provide an stable activation path.

b. Interpretation of MSOM learning

From the description of the MSOM algorithm in the previous section, MSOM learning converges toward the representation of the input patterns, expressed through the weight vectors. For input slices corresponding to the same time t , the network activation should occur at different neuron locations if the time slice is different. Similarly, time slices corresponding to input patterns with different time characteristics (e.g. different decay rates spectrum) should activate paths of different lengths and positions. In principle, one should expect that a pattern can be identified by the *bmu* corresponding to the first or last time slices or by the path of *bmu*'s.

One possible interpretation of the MSOM learning is the formation of clusters of *bmu* paths that defines non-overlapping or overlapping areas in the map. In the first case, a trained MSOM can be used to identify input patterns by just assigning a class to each neuron after *a-posteriori* analysis of the paths, like in the trained MSOM cartoon depicted in the bottom right of Figure 28. In the second case, since parts of the activation paths may share the same or nearby neurons, the association of an input pattern to a specific cluster should be done in several different ways: i) by assigning the class of the closest stored path using a least-squares approach or ii) by defining an extra FFANN or SOM layer to provide a non-linear classifier or clustering stage. In section 3, we will continue the discussion about MSOM output interpretation.

D. Datasets and pre-processing

Training data can be organized in two groups: static and spatiotemporal input patterns for analysis using SOM and MSOM, respectively. Each group contains both synthetic and real data. Synthetic data provided a way to verify that the algorithm was performed as expected, learn about the appropriate range for SOM and MSOM parameters, neighborhood initial size and rate of shrinking, learning rates, neural network size as a function of the dimension and size of the dataset, computation time, etc.

For static synthetic input patterns the datasets are the following; i) two concentric, overlapping Gaussian distributions (dimension 3), ii-iii) transient signals with two and three clusters with different decay rates distributions (dimension 26), and iv) spatial profiles from four different magnetic dipole sources (dimension 25). For synthetic spatiotemporal input patterns, two datasets were generated consisting of time-dependent dipole responses with two and three different exponential decay rate distributions (dimension 13×26). In the case of experimental datasets, static feature vectors of dimension 4 representing three different classes corresponding to the Iris dataset, central loop receiver CSEM transients (dimension 26) and spatiotemporal profiles using a modified Geonics EM63 (11 recording station positions with 26 timechannels each, see Appendix A) were also analyzed.

a. Benchmark datasets for SOM algorithm

Synthetic transients and spatiotemporal datasets were generated trying to keep experimental setup and background-subtracted signal parameters of the real datasets obtained with the Geonics EM63. Even though the way the data was calculated corresponds to very specific and ideal conditions, their parameters are very close to the

real datasets. The main difference is the lack of random noise always present in real data mainly due to electronics and environmental noise.

The first dataset consist of two clusters (40 vectors each, dimension 3) of randomly generated Gaussians realizations, centered at the origin with variances σ 0.4 and 0.8. Density of data is larger for the distribution with smaller variance sharing the space with several points of the broader, less dense distribution (see Figure 29).

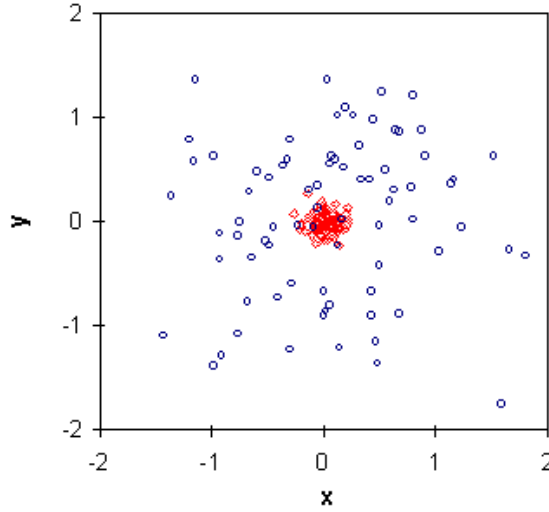


Fig. 29. Two concentric Gaussian clusters dataset.

Gaussian distributions are followed by two datasets comprising decay signals. The first one consists of two families of 40 transient waveforms: single exponential decays ($A_1 \exp(-s_1 t)$), and stretched exponential decays ($A_2 t^{-\beta} \exp(-(s_2 t)^{1-\beta})$). The second includes an extra cluster; 40 two-exponential cluster of decays ($A_3 \exp(-s_3 t) + A_4 \exp(-s_4 t)$).

Variability in both groups was produced by perturbing A_i , s_i and β with realizations of white noise with maximum amplitudes that were 1/10 of the reference value. Reference values for the first dataset are $A_{1,2} = 1$; $s_{1,2} = 700, 300$; and $\beta = 0.6$ while for the second one $A_{1,2,3,4} = 1, 0.8, 0.3, 0.002$; $s_{1,2,3} = 700, 700, 1500$; and $\beta = 0.7$.

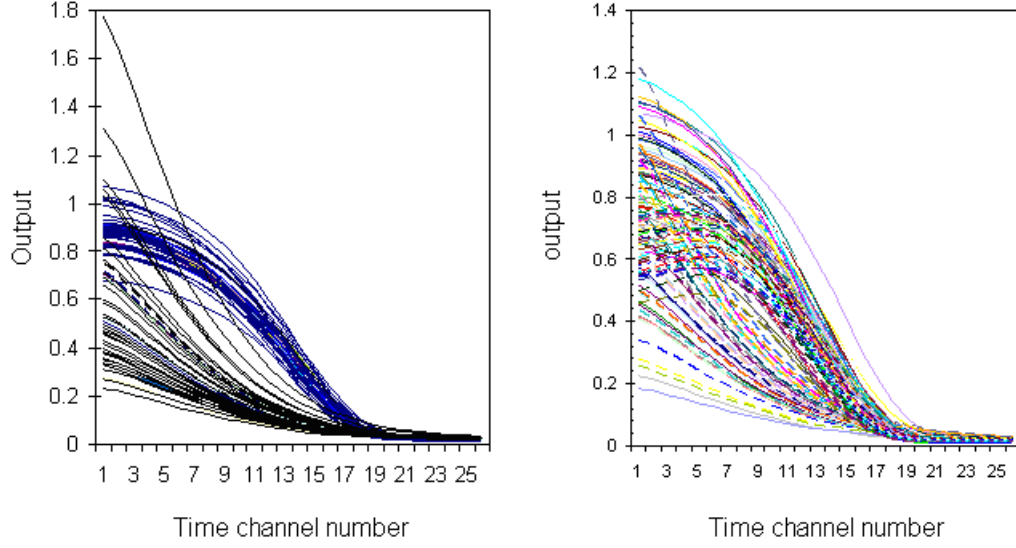


Fig. 30. Datasets from synthetic transient data. Left: 2-cluster data; single exponential and stretched exponential decays. Right: 3-cluster data; one exponential, stretched, and two-exponential decays.

EM63 timechannels values were used to generate the dataset. Figure 30 shows the amplitude of the transients as a function of the timechannel number, which is equivalent to a plot in a logarithmic scale of time.

Dataset consists of four different 1-D dipole responses (dimension 13) representing possible EM63 profiles at specific timechannels for dipole sources with different orientations (declination and inclination). For each cluster, parameters describing unit magnetic moments located to a depth of $z_o = 0.15$ m, are as follows: first cluster corresponds to horizontal dipole parallel to the profile direction (class 1), the second cluster are vertical dipoles (class 10), the third cluster are dipoles inclined 45 degrees in the profile direction (class 20) and the last cluster is generated by dipoles oriented horizontally and perpendicular to the profile direction (class 30). Each cluster consists of 20 profiles.

In order to generate each cluster, parameters were perturbed using the same procedure described with synthetic transient datasets but in this case, a Gaussian

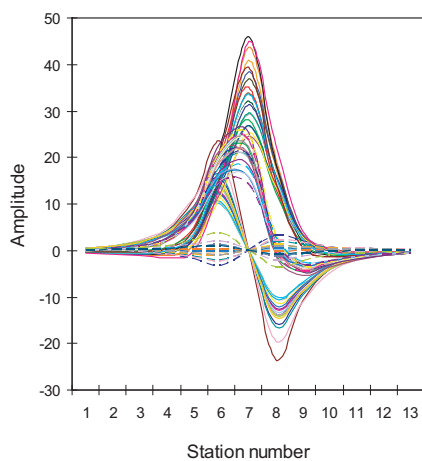


Fig. 31. Dataset of spatial dipole responses.

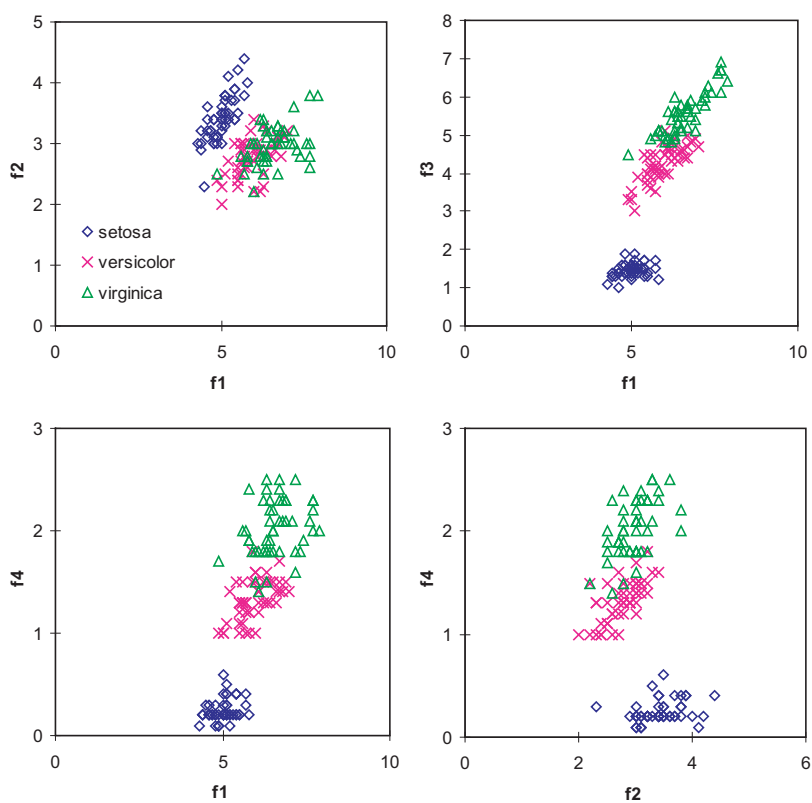


Fig. 32. Some data projections for the Iris dataset.

PDF with variance 0.01 was used. Figure 31 shows the whole dataset. There are clear differences between classes in terms of waveforms and amplitudes.

Iris Dataset is likely the best known dataset¹ found in pattern recognition literature. The dataset consists of 50 input vectors for each one of the 3 classes. Class refers to a type of iris plant namely Setosa (0), Versicolor (0.5), Virginica (1.0). Each input vector contains 4 numeric attributes or features (f_i). The first class is linearly separable from the others while the latter are not linearly separable. Figure 32 shows the dataset as 2-D projections.

b. Benchmark datasets for MSOM algorithm

For MSOM algorithm testing, two spatiotemporal datasets comprising two and three clusters (10 patterns each) of dipole responses decaying in time were generated. The dimension of each input pattern was 13×26 . For the first dataset, the parameters were the following: $m_1 = 1$; $z_{rx}(1, 2) = 0.25$ m; $z_o(1, 2) = 0.15$ m; decay rate $s_1 = 700ms^{-1}$; $D(1, 2) = 0$; $I_1 = 0$; $m_2 = 0.8$; $s_2 = 900ms^{-1}$; $I_2 = \pi/2$.

Figure 33 shows each cluster of patterns in separate plots. Each plot is built by appending the 26 spatial responses (time slices) one after another, following the right time sequence. A vertical grid of lines helps to separate each spatial profile. The amplitudes in the graph are square-root compressed and normalized. Differences in the time can be noticed in the last six spatial responses. There is a clear difference between classes due to the spatial waveforms. In principle the dataset should be separable if the MSOM algorithm performs correctly.

In the second dataset, there are three clusters: two decaying exponentially in time and one decaying with stretch-exponential decay. The parameters were the

¹Available at the URL: <http://einstein.library.emory.edu/PRS-data.html>

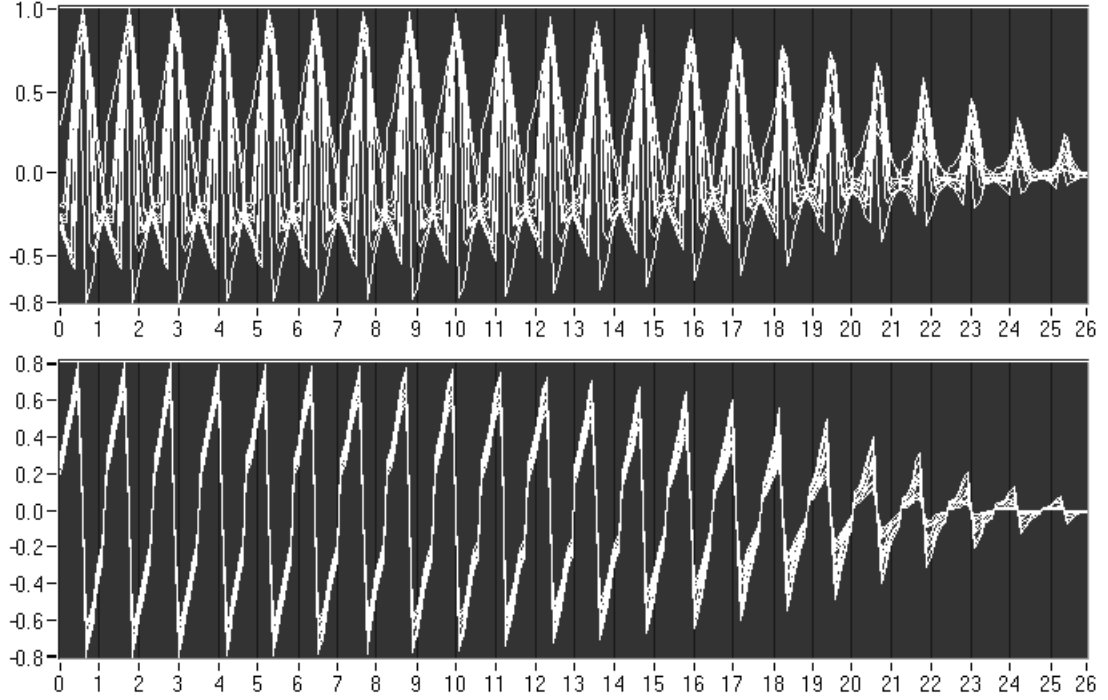


Fig. 33. Two-cluster synthetic spatiotemporal data.

following: $m(1, 2, 3) = 1.0$; $z_{rx}(1, 2) = 0.25$ m; $z_o(1, 2) = 0.15$ m; $s(1, 2) = 700s^{-1}$; $D(1, 2) = 0$; $I_1 = 0$; $I(2, 3) = \pi/2$; $\beta = 0.4$; $s_3 = 400 s^{-1}$. In order to generate each cluster, parameters were perturbed using Gaussian realizations with variance 1.0 and amplitude $0.2m_i$. In this case (Figure 34), classes 2 and 3 are spatially similar, with an important temporal differentiation at late time. Class 1 is spatially more complex but its time decay is in between classes 2 and three, meaning that clustering this dataset would require an efficient context manipulation by the MSOM algorithm.

c. Geonics EM63 data

Data were obtained using a Geonics EM63 provides a high-quality commercial version of a time-domain metal detector. With a 16-bits analog to digital converter and real time stacking, the receiver unit can record transient data from large conductive targets

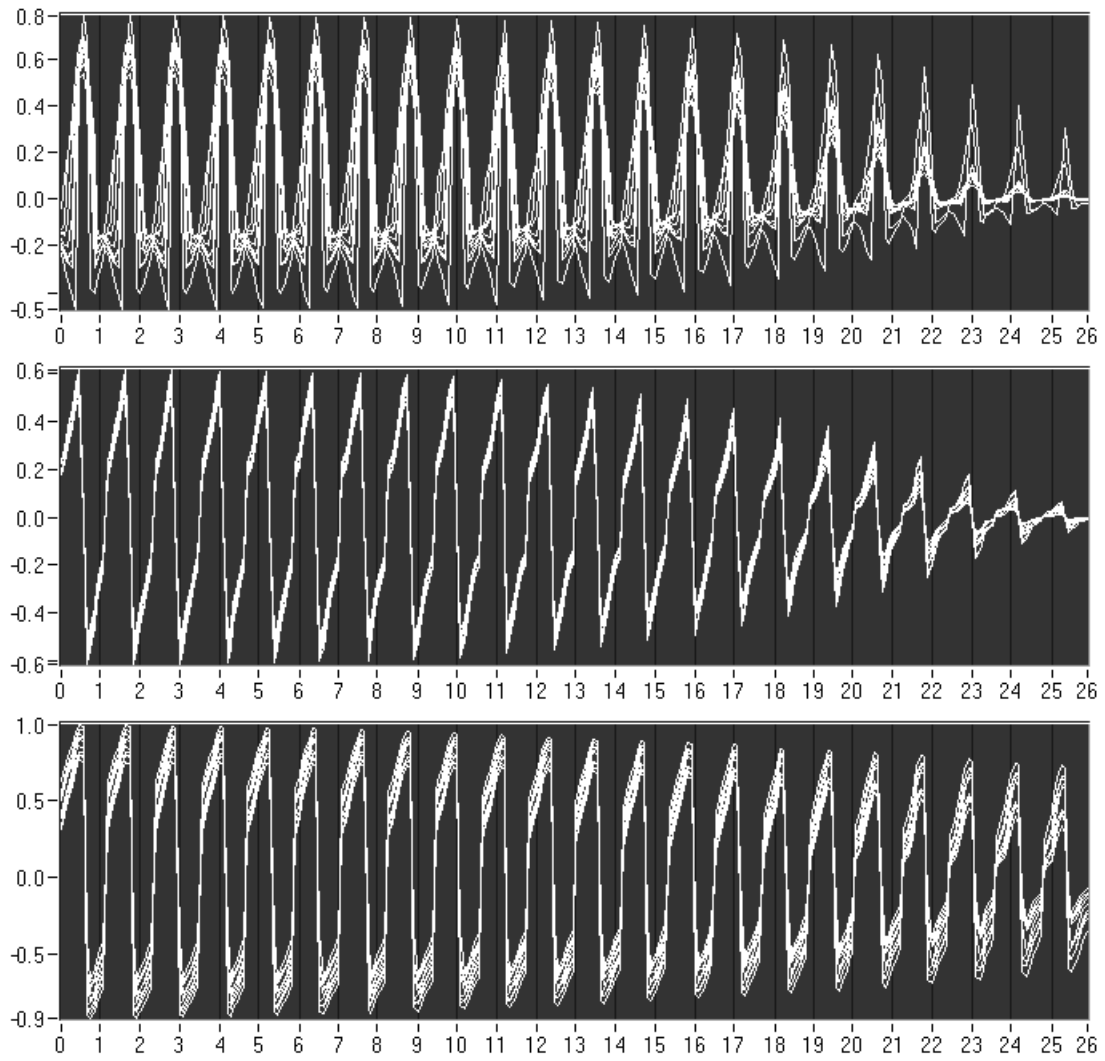


Fig. 34. Three-cluster synthetic spatiotemporal data. From top to bottom; classes 1, 2, and 3.

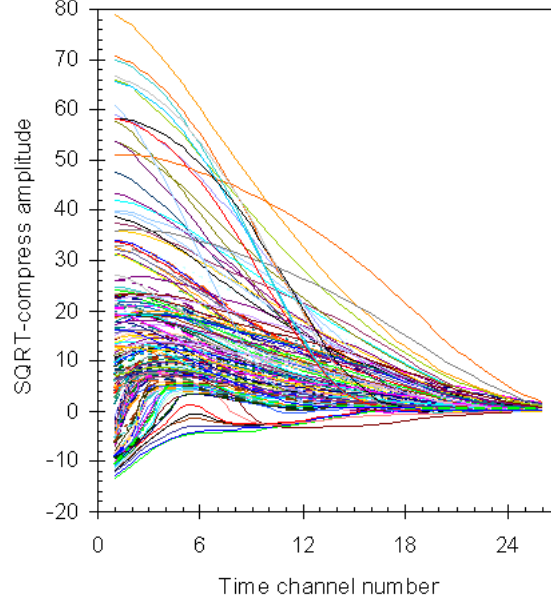


Fig. 35. EM63 square-root compress transient dataset.

from $10V$ down to the noise floor, which is about $20\mu V$. In this work, only the vertical component of the secondary, time varying magnetic field was acquired.

Central loop transients were acquired with the receiver coil array occupying the center of the Tx coil, $0.25m$ above the ground surface as showed in Figure 19. Spatiotemporal profiles consisting of 25 Rx positions, as described on section III-C, were used for testing clustering capabilities of a MSOM. Since the total profile length ($2.5m$) and Rx interval ($0.1m$) is large and dense, respectively, for a potential cart-mounted multi-receiver instrument, a shorter, $2.0m$ profile with $0.2m$ of Rx interval was extracted from the data, providing a 26×11 matrix of receiver output voltages for each spatiotemporal pattern. This data decimation still preserves general features of the time slices, simulates a more realistic instrument dataset and reduces the time required to train a MSOM. About half of the central loop transients used in the static vector dataset were extracted from the central Rx location in the spatiotemporal dataset. The remaining 50% was acquired during numerous field tests involving ad-

ditional background and target measurements using the EM63 as a modified, central loop system. The complete transient dataset is shown in Figure 35.

Different kinds of targets were buried at five different locations on Texas A&M University Riverside Campus. Depth to the target top was varied between 0.1 and 0.45 m. In the case of axi-symmetric targets, data acquisition was repeated with the target buried in seven different orientations with respect to the ground surface and profile direction, as shown in Figure 36. Compared with the previous datasets, the EM63 spatiotemporal patterns show important diversity of waveforms and time decays, especially when comparing the patterns from ordnances and clutter. Central loop data consist of 201 transients distributed as follows: 42 from the subsurface without any target present (BG), 114 containing both subsurface and ordnance responses: 38 in orientation C1, 12 in C2 and C3, 17 in C4, 13 in C5, 11 in C6 and C7, and 47 from several non-ordnance targets (clutter) such as steel wedges, spheres and discs; aluminum plates, spheres, soda cans and disks; closed copper loops. Spatiotemporal data contains 130 patterns: 12 from BG, 78 from ordnances and 40 from clutter items. The whole dataset is plotted in Figure 37.

A large dynamic range in the input patterns is inconvenient for training artificial neural applications because the weights adaptation will pay more attention to the elements whose amplitude is larger. At the same time, amplitude information should be preserved somehow in order to allow separation of different input patterns produced by targets with large differences in size or electromagnetic coupling with the instrument's transmitter and receiver. In order to reduce the dynamic range of the transients, and at the same time, to preserve and enhance temporal differences that occur at late time, a square root or logarithmic amplitude compression was used;

$$x_{sqrtc} = sign(x)\sqrt{|x|}, \quad x_{logc} = sign(x)\ln(|x| + 1). \quad (4.23)$$

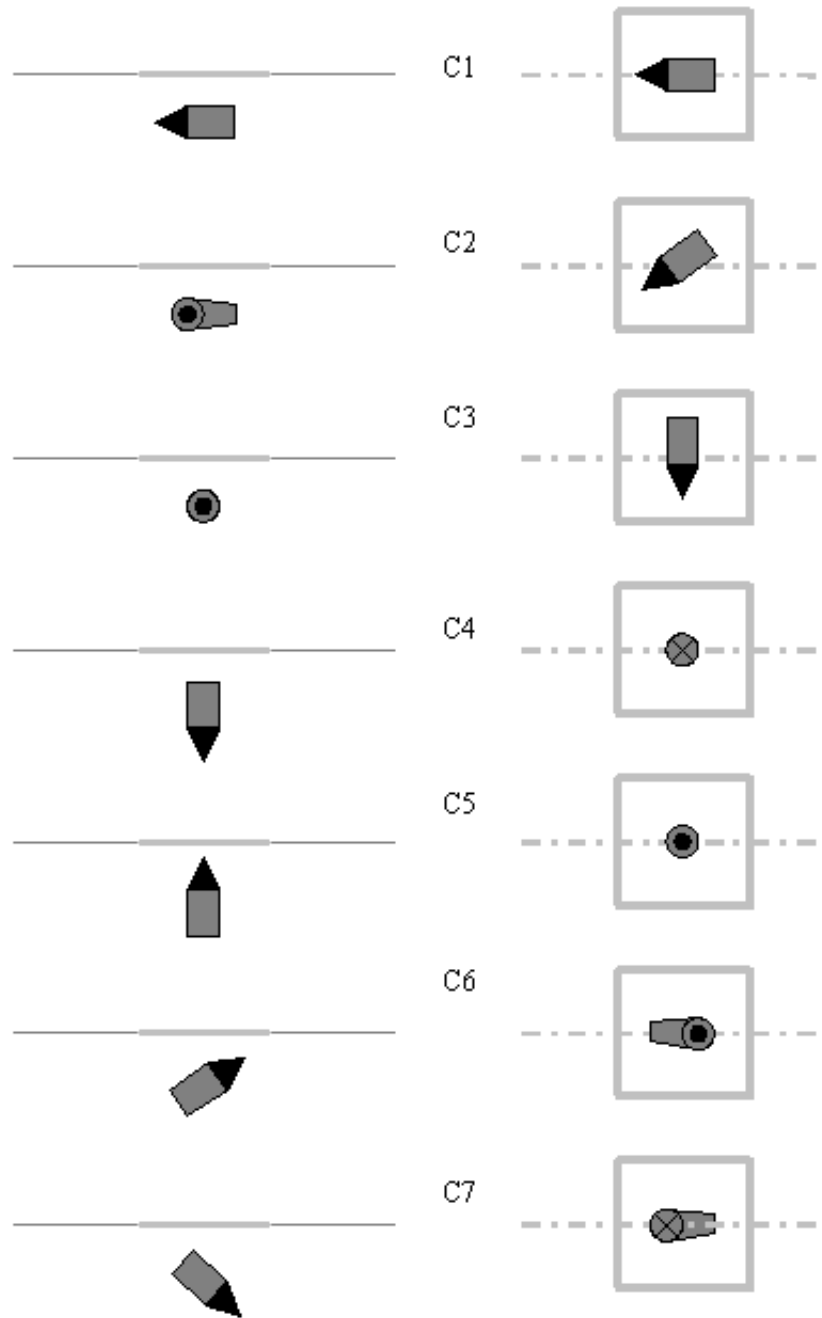


Fig. 36. Target burial orientations. Side (left) and map views (right) of the seven different target configurations used on the EM63 synthetic multi-receiver experiments. Station 0 lies on the left of the profile line.

Scaling of the data to provide the same statistical profile $(0, 1)$ was also applied to synthetic data other than transients. In some tests with spatiotemporal patterns, an additional scaling $\mathbf{X}' = \mathbf{X}/(\max \mathbf{X})$ was applied to the whole dataset to assess improvements in learning time.

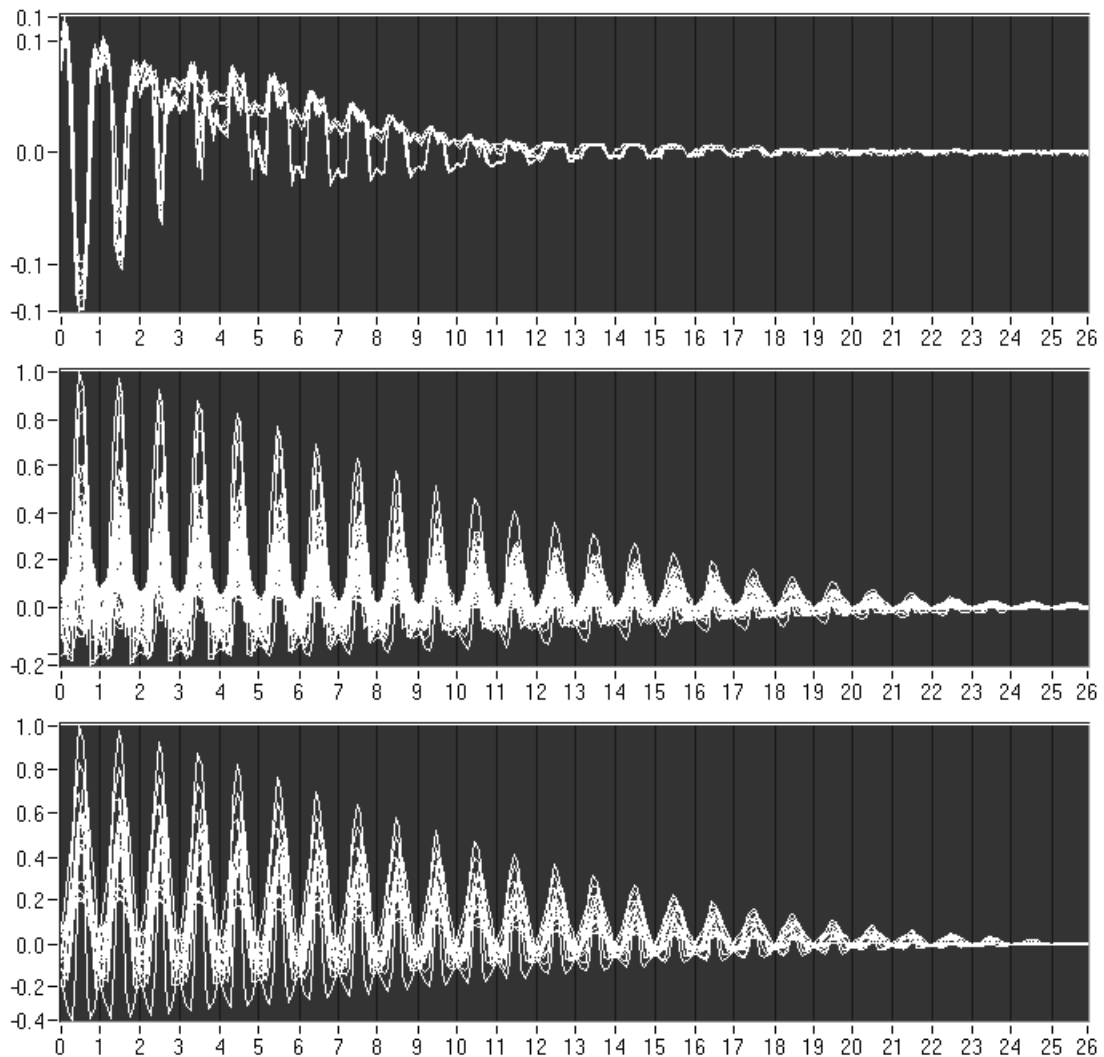


Fig. 37. Geonics EM63 spatiotemporal responses. The dataset is shown as three categories: pure subsurface or background (top), unexploded ordnance (middle), and clutter items (bottom). Amplitudes as plotted with square-root compression and dataset normalized.

3. Results and discussion

The results are organized in two main groups: datasets used for algorithm benchmarking and modified Geonics EM63 CSEM datasets. The latter consists of two groups: central loops transient data comprising the vertical component of the secondary magnetic field, and spatiotemporal datasets obtained from 1-D profiles of transients above the location where different kinds of conductive targets were buried at depths ranging from 0.1 to 0.4 m.

A. Benchmarking of SOM and MSOM algorithms

Results are presented by providing the algorithm settings, a plot visualizing the resulting clusters, using U^* -matrix representation, and the network activation produced by presenting the training dataset to the trained network. In the latter, each active neuron will appear as a pixel with color representing the class the input pattern belongs to. The results follows the following order: two Gaussians, central loop transients, dipole response patterns, and the Iris dataset which represents static input patterns processes with SOM. The section ends by presenting the results from the synthetic spatiotemporal datasets, processed with MSOM.

For the two concentric Gaussian distributions, a rectangular SOM with 400 neurons was trained with learning rate of 0.8 and an initial neighborhood width of 18. Training took 500 epochs to converge. A Pareto radius of 0.025 was used for the P -matrix calculation. The U^* -matrix in Figure 38 shows a central “valley” surrounded by an irregular boundary. Testing of the input data in the trained network (Figure 38, right) shows a perfect separation of both clusters; vectors belonging to the narrow Gaussian distribution fell in the central valley while the others were scattered inside small plateaus outside the boundary. A complementary test using K-means

and FFANN algorithms failed to classify the data correctly. For the first synthetic

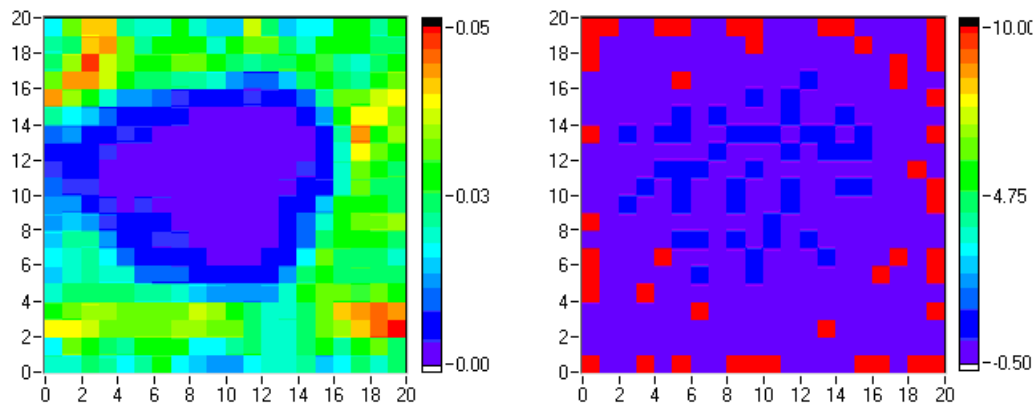


Fig. 38. SOM results for two concentric Gaussian distributions data. U^* -matrix (left) and the map activation from presenting the training dataset to the trained network (right).

transient dataset, a 15×15 SOM was trained using a linearly decreasing learning rate, from 0.9 to 0.0004. The neighborhood function had an initial width of 18, decreasing exponentially. Convergence was attained after 300 epochs. The U^* -matrix, calculated using a Pareto radius of 0.23, shows a very well defined cluster boundary (Figure 39) and perfect separation of the two classes.

For the second synthetic transient dataset (Figure 40) a 20×20 SOM was trained using the same learning rates and neighborhood function. The training converged after 900 epochs. Using a Pareto radius of 0.18. The U^* -matrix shows a clear three clusters structure. The test of the input pattern verifies that the class assignment is 100% correct. Notice that the network organizes itself so it spreads the different input vectors over the whole map.

In the case of dipole response patterns, a 20×20 SOM was used for this dataset. Setup parameters were: Initial learning rate = 0.8; Learning rate decay type = linear; Minimum learning rate = $1.0\text{E-}4$; Neighborhood kernel = Gaussian; Initial variance = 17; Variance decay type = exponential; Variance decay rate = $5.0\text{E-}3$; Number

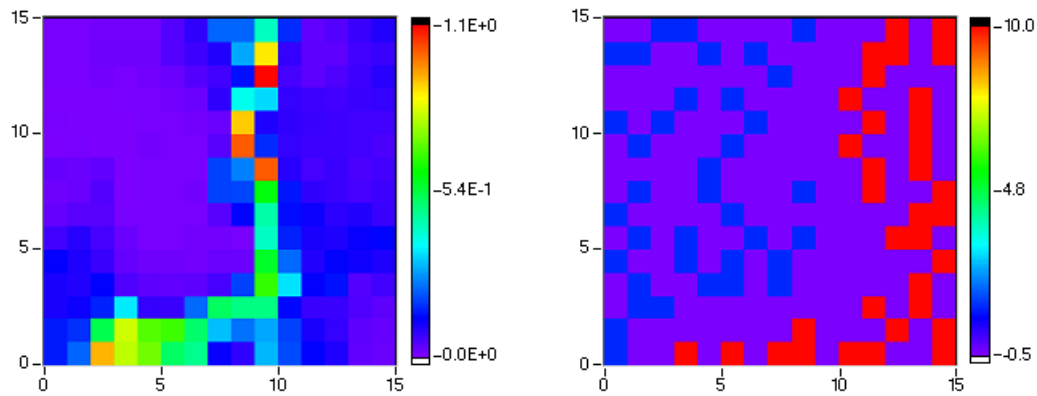


Fig. 39. SOM results for 2-cluster of transient data. U^* -matrix (left) and the resulting map activation from presenting the training dataset to the trained network (right).

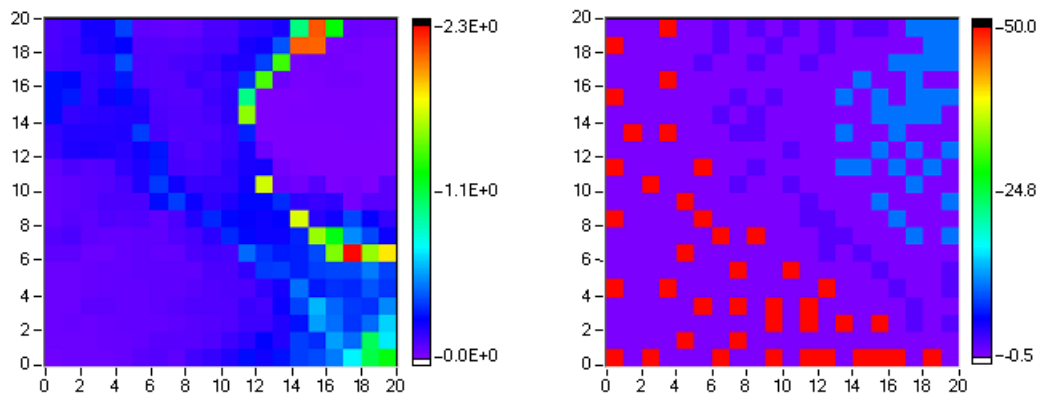


Fig. 40. SOM results for 3-cluster of transient data. U^* matrix and the resulting map activation from presenting the training dataset to the trained network (right).

of epochs = 1,000. The U^* -matrix was calculated using a Pareto radius of 7.78, corresponding to the percentile 18%. Figure 31 shows the dataset. Notice that this dataset should be easy to map since the differences between clusters are more or less obvious.

Figure 41 shows that the U^* -matrix has produced a sharp and somewhat tortuous class boundary that basically creates four cluster areas. Only class 10 belonging to the vertical dipole sources seems to have a sub-cluster in the bottom right corner of

the U^* -matrix. Nowadays, checking the input test shows that the mapping is correct since there is no pattern projected in the wrong U^* -matrix cluster.

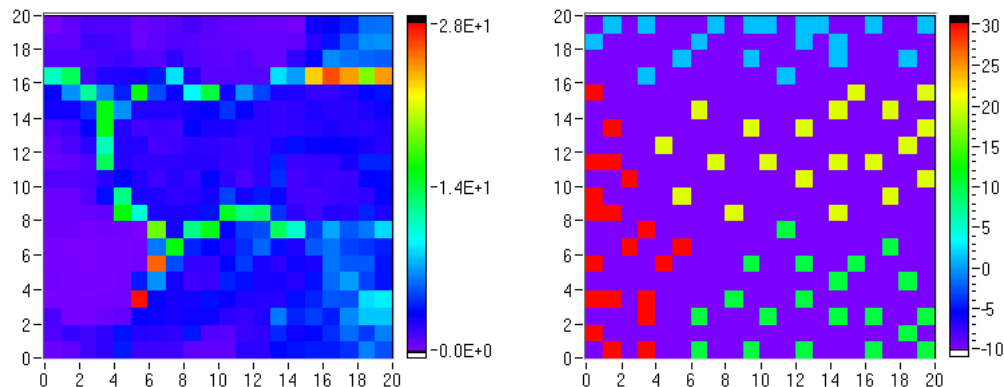


Fig. 41. SOM results for 4-cluster of dipole responses. U^* matrix (left), and activation map resulting map activation from presenting the training dataset to the trained network (right).

The Iris dataset completes the group of benchmarking static datasets. In this example, a SOM with 20 x 20 cells was used. A learning rate ($\gamma = 0.8$) and a Gaussian Neighborhood function ($\sigma = 18$) decreasing exponentially was used. The input vectors were normalized and randomized. Training took 400 epochs to reach the stopping condition. Top of Figure 42 shows the resulting U^* -matrix, using a Pareto radius of 0.2, in which 4 different “valleys” separated by curved ridges of different height. The highest boundary separates class 0 which is linearly separable from the other two. Class 0.5 corresponds to the central, hourglass shape cluster that seems to be splitted in two in the U^* -matrix representation. The lower boundary separates class 0.5 from class 1. By presenting the input dataset to the trained SOM and assigning class number to color intensities, a map of *bm*u can be easily obtained (see Figure 42, bottom. Notice that the network misclassify two class 0.5 inputs as belonging to class 1. The same result was obtained with supervised learning using K-means and

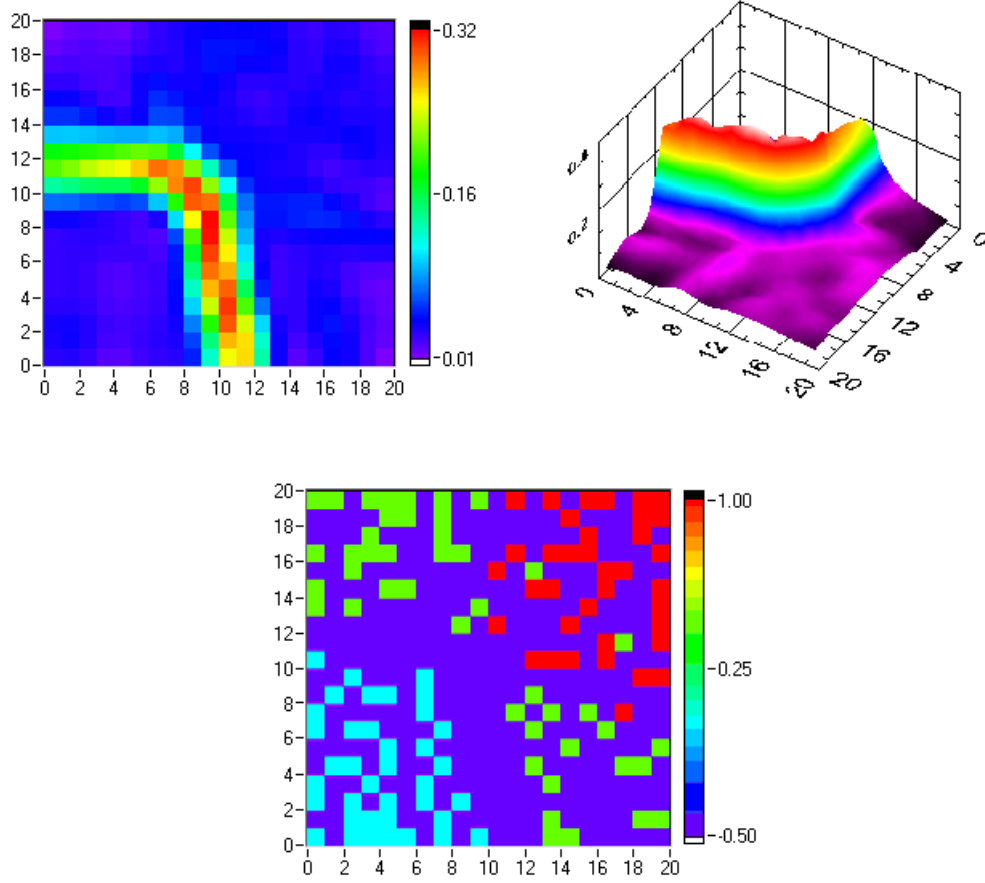


Fig. 42. U^* matrix and map activation for the Iris dataset. Visualization of the U^* matrix for the Iris dataset (top) as a pixel map and as a rendered 3-D map. SOM activation by the training dataset (bottom), shows pixels in blue, red and green representing bmu for classes 0, 1 and 0.5, respectively.

FFANN with seven neurons in layer one, 6 in layer 2 and one output neuron.

To finalize this section, MSOM results for two spatiotemporal datasets are presented. MSOM setup parameters that succeeded in training the network for two cluster spatiotemporal dataset were: initial learning rates $\gamma_{w,c} = (0.12, 0.12)$ with a linear decay down to 10^{-6} . A Gaussian neighborhood with initial variance $\sigma = 18$, and exponential decay rate of 5.0×10^{-3} produced good convergence results in a 20×20 MSOM network after 1,000 epochs (1.6 hours of training). Weights and context were stored on each neuron as a 26-dimensional tuple. Network dynamics

was controlled using a merge parameter α varying linearly in the range 0.01-0.36. The merge context parameter was kept constant ($\beta = 0.5$). Dataset was square-root compressed and randomize prior to the training phase.

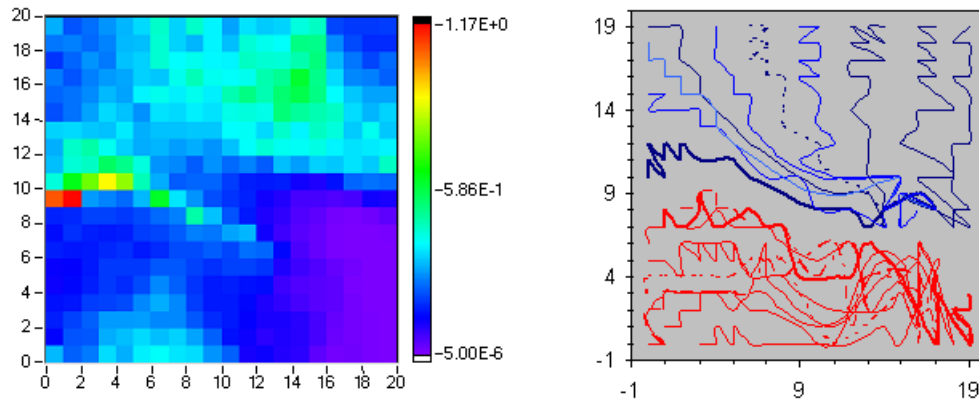


Fig. 43. MSOM results for 2-cluster synthetic spatiotemporal data. U^* -matrix (left) and $bm\mu$ paths (right).

Figure 43 shows the results of the training. The U^* -matrix shows a cluster boundary in the bottom third of the map that produces a neat class separation. However, without a post-training plot of the $bm\mu$ paths, it could have been difficult to dismiss two “topographic highs” in the upper left and bottom right of the U^* map. Notice that learning has spread the different input patterns in such way that the first time slices have been located at the top, left and part of the bottom edges of the network, paths were arranged running in an ordered fashion to the bottom right corner. Set of input patterns with decay rates about $900ms^{-1}$ occupies the bottom valley (red paths), while the slower decaying patterns took the upper two-thirds of the map. A close look to both classes final $bm\mu$ ’s shows that even though they reached a valley in the U^* -matrix map, they activated different groups of neurons, meaning that in this simple two-class example, clustering or classification of the input patterns can be achieved by observing the position of the first or last $bm\mu$, or by defining two

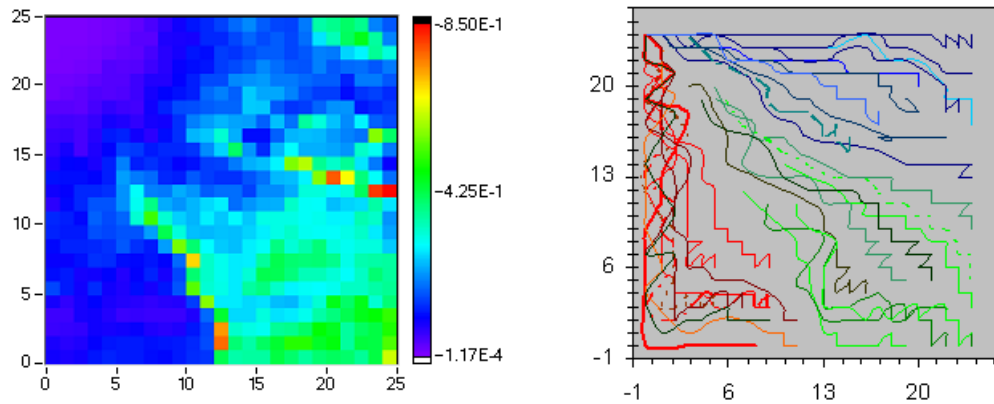


Fig. 44. MSOM results for 3-cluster synthetic spatiotemporal data. U^* -matrix (left) and bmu paths (right).

different clusters above and below the U^* -matrix “ridge”. Clearly, the MSOM was efficient in using context for discriminating two different types temporal content.

For three cluster spatiotemporal dataset, MSOM setup parameters were basically the same as in the previous example with exception of the initial learning rates $\gamma_{w,c} = (0.1, 0.1)$, initial variance and decay rate of the neighborhood ($\sigma = 20, 4.0 \times 10^{-3}$). The training required 1.9 hours of training to be completed. Weights and context were stored on each neuron as a 26-dimensional tuple. Networks dynamics and merge context were $\alpha = 0.4$ and $\beta = 0.4$, respectively. The results, depicted in Figure 44 shows the same general characteristics described in the previous paragraphs but now, two cluster boundaries separate three “valleys” in the U^* -matrix map. Red paths in the bottom most valley corresponds to the first cluster (with parameters about the values $s = 700ms^{-1}$, $D = I = 0$), paths in green in the central valley are the projections of the patterns about (parameters about the values $s = 700ms^{-1}$, $D = 0$, and $I = \pi/2$), and the top most cluster contains the stretch exponential decays.

B. SOM for EM63 central-loop transients

The dataset with 203 EM63 central loop transients was the training target of a 40×40 SOM. An initial learning rate γ of 0.9 decaying linearly down to 10^{-5} along with a Gaussian neighborhood with initial variance $\sigma = 40$ that decreased exponentially (decay rate 4×10^{-3}) provided the best convergence. A total of 2.3 hrs of learning were necessary to complete the training phase (203,000 input patterns presentations).

The U^* -matrix was calculated using a Pareto radius of 0.18. For a *posteriori* class labeling, the dataset was labeled into seven classes and presented to the trained network. Labels for background responses (only ground response 10, noisy background 50), UXO (90), and clutter (Cu loops 130, plates 170, spheres 210, and other types of clutter; disks, wedges, cans, 250) were used to define cluster regions, by comparing the activated neurons with the boundaries enhanced by the U^* -matrix. Cluster boundaries were chosen along “highs” or ridges in the U^* -matrix.

Figure 45 shows the resulting U^* -matrix, input data activation map of the trained network and a possible definition of the cluster boundaries that can be used to assess the label of a new input pattern. The U^* -matrix shows a complex pattern of ridges separating areas of diverse size; background patterns are very well clustered in the right side of the map, continuous, high boundaries. The largest map area is covered by UXO, which represents 56% of the dataset. Input patterns from some coils, soda can, small aluminum disk, plates, and solid spheres were activated in small regions at the edge of the network, separated from the UXO region. The area covered by UXO is not uniform, meaning that patterns from other types of targets were mixed with UXO, forming small groups of one or few activations. For some of those mislabeled patterns the U^* -matrix produced a funnel which isolates that particular pattern inside the UXO “cluster”. Eight clutter patterns from spheres and plates were mislabeled

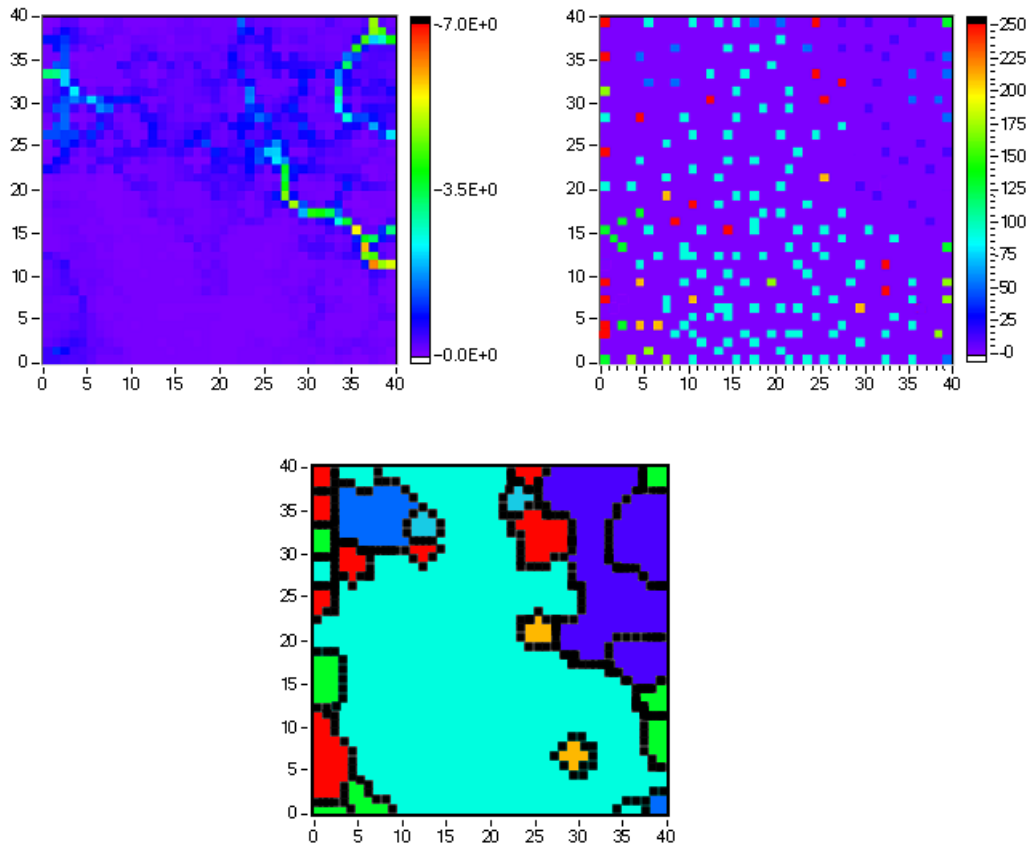


Fig. 45. SOM results for EM63 central-loop transient data. The U^* -matrix (top left) along with the input data activation map (top right) was used to define an interpretation map to assess the label of a new input pattern (bottom) using the same color code of the activation map.

in a *a posteriori* labeling which included both input and testing input patterns, the latter representing 10% of the input dataset. This can be translated in a 16% of misclassification of clutter as UXO. Background patterns were successfully identified, providing 100% classification capability. The misclassified targets were two hollow iron spheres, $0.3 \times 0.3m$ aluminum plates and UXO parts.

Despite the fact the dataset is not large or balanced enough to assure an equal partitioning of the network map among the classes it is quite clear that the SOM method can potentially provide a good interpretational instrument to classify target responses which includes the response from the ground materials.

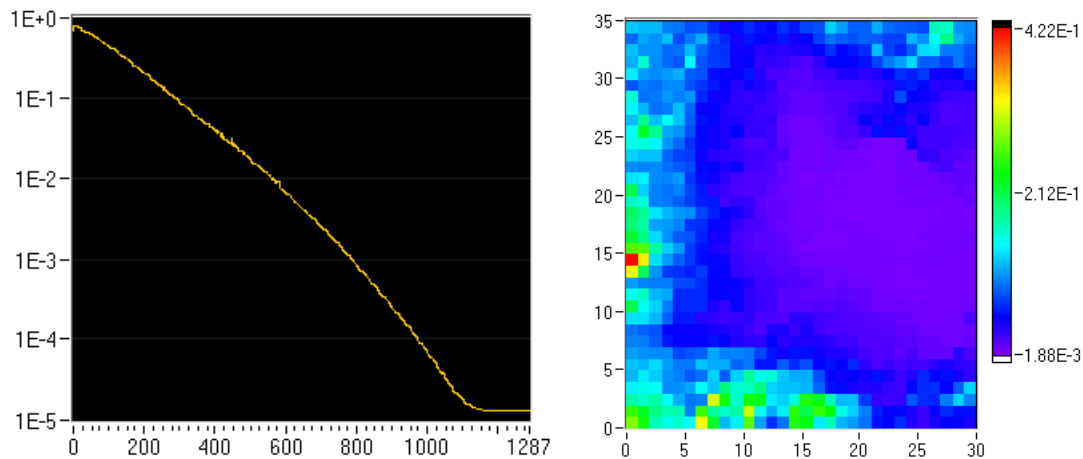


Fig. 46. Average cost function per epoch and U^* -matrix for EM63 spatiotemporal data.

C. MSOM for EM63 spatiotemporal patterns

To conclude the results section, spatiotemporal EM63 multi-receiver data (see Figure 37) analysis using a MSOM is presented next. A 35×30 MSOM was trained using a fixed, 0.1 learning rate for both the weight and context learning rule. The dynamical parameter α was chosen to start with a value of 10^{-2} and was increased linearly to up 90% of the training phase, after which it was kept constant with a value of 0.2. The merge parameter β was fixed at 0.5 during the whole training while a Gaussian neighborhood with initial variance $\sigma = 29$ decaying exponentially with rate 3.8×10^{-3} defined the range and duration of the cooperation between neurons. Each neuron stored a tuple (\mathbf{w}, \mathbf{c}) of dimension 22. The dataset was square-root compressed and scaled such that the dataset was in the interval $(-1, 1)$, the sequence of patterns randomized. During the testing phases of the MSOM for processing EM63 spatiotemporal data, numerous network sizes (from 20×20 up to 40×40) and parameter values were tested. It was verified that high values (close to 0.5) of the dynamical parameter α produced not good results in the pattern learning nor in the convergence rate. Moreover, a variable α , starting with small values as

the one indicated above, produced better clustering of the paths and improved the reduction of the average cost function per epoch, especially during the early stages of the training. Training phase ranged from 6.8 to 10.1 hours to converge (using a PC with a x86 Intel Centrino processor and 1.28 GB of memory). For the MSOM to be shown next the training time was 9.49 hours for 1,286 epochs. Figure 46 shows that the training process reduced the average cost function (per epoch and neuron) from almost 10^0 to 10^{-5} with a clear stabilization shortly before the training reached 1,200 epochs. Figure 46 also shows that the resulting U^* -matrix presents a dominant low

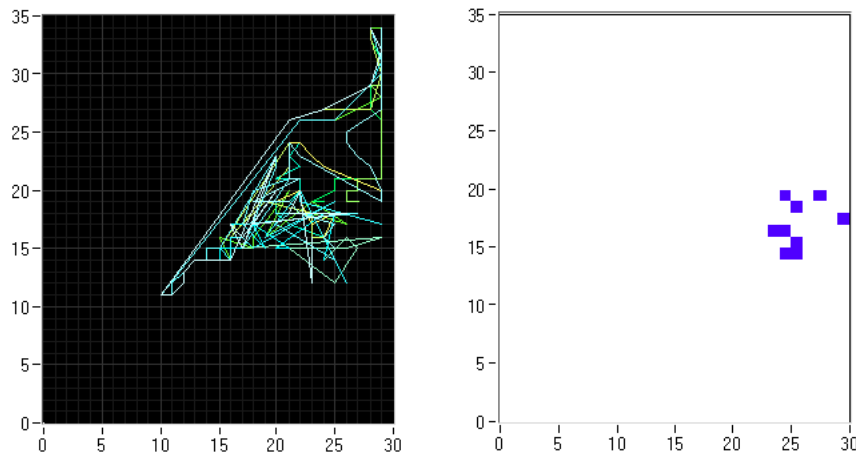


Fig. 47. MSOM activation paths for the EM63 spatiotemporal background data. Background paths (left) and their respective last *bm_u* (right).

in the center right surrounded by variable heights at the top, right and left, providing that there are multiple paths starting from those three edges and converging in the right side of the map. In order to demonstrate how the MSOM learn to project the input patterns, a series of plots will follow starting with the background responses Figure 47 that shows the sequence of *bm_u*'s for each input pattern as a path in the network space. Also, the activated neuron for the last time slice is shown alongside the paths. Notice how the background paths are limited to a well defined area of the

map, starting in the top right corner and finishing halfway down in the same map side.

In contrast, the paths for several UXO input patterns are very different. Figure 48 shows that for UXO in orientation C3 (ordnance horizontal and perpendicular to the profile direction), paths start at the top and left edges of the network, ending around the same area as the background paths, similar to what we observed with the synthetic spatiotemporal datasets. Similar paths are observed with the UXO patterns in orientation C5. Comparing background and UXO paths with clutter

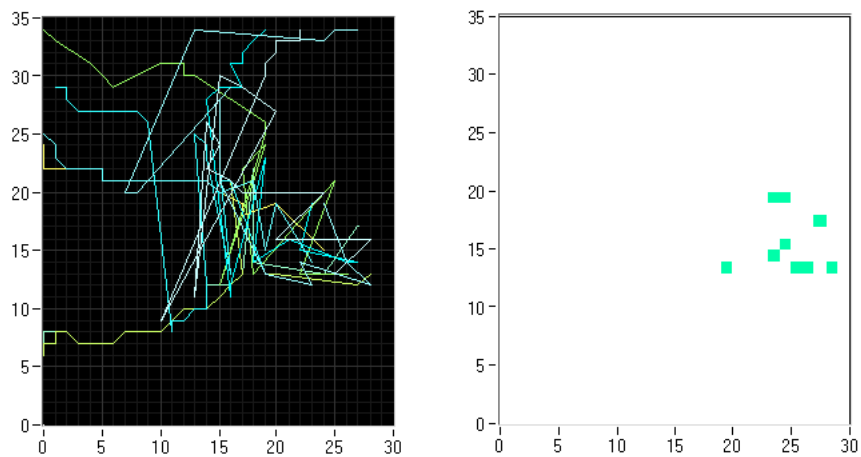


Fig. 48. MSOM activation paths for EM63 spatiotemporal UXO data in orientation C3. Paths (left) and their respective last *bmu* (right).

paths can provide a good assessment about the capacity of a MSOM to discriminate between these three types of sources. Figures 49 and 50 shows clutter paths when the items have been buried horizontally or vertically. Notice that path geometry in these cases resembles background paths. Similar results have been found for clutter of axis-symmetric targets in configuration C6, plates, loops and spheres, improving the discrimination capability with respect to SOM for central-loop transients. So far, it has been shown that MSOM makes an appropriate use of the network space

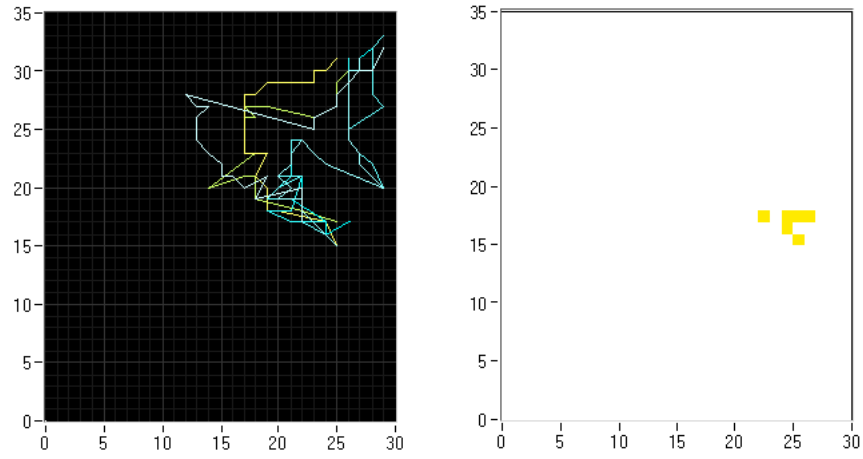


Fig. 49. MSOM activation paths for EM63 spatiotemporal clutter data in orientation C1. Paths (left) and their respective last *bmu* (right).

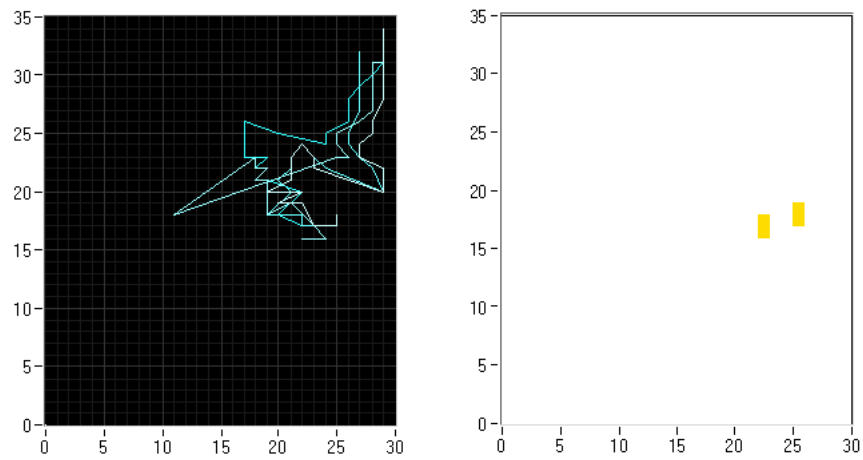


Fig. 50. MSOM activation paths for EM63 spatiotemporal clutter data in orientation C4. Paths (left) and their respective last *bmu* (right).

and neuron context projecting input patterns according to its spatial and temporal characteristics into paths with different geometry, length, initial and final *bmu*.

One valid question is: how good is the map activation after presenting an input pattern? If MSOM is really learning about the input patterns, it should be able to produce accurate spatiotemporal responses. At least we should expect that if we present training input patterns. It turns out that an analysis of the MSOM activation produces activation patterns with a *rms* error in the range of 0.0028–0.0084 for all the input patterns. Figure 51 shows the input pattern as blue squares and the activated pattern as a red line. In order to keep track of the quality of fitness for points corresponding to the late-time slices, a plot of the activated values as a function of the input values is shown, having a slope-one line as reference. Figure 52 shows that there is more dispersion in the paths but still they are different from the background paths.

Another example on the same subject is shown for clutter in orientation C1 in Figure 53. Clearly, MSOM has learnt important details about the input patterns and its activation is very easy; it is only necessary to present the input pattern using the final dynamic parameter α and merge parameter β .

A trained MSOM can be used for pattern classification by defining the neurons most visited for each type of input patterns. If that is the case, the network can be labeled in a similar way to what it was done in case of SOM for transient data. Two more general approaches that does not rely on operator criteria is to store the paths as a complementary library and after the map produced an activation path, it is compared with the path library to identify the one with smaller *rms* error. Another option could be to add a trained FFANN layer in which input of the network neuron coordinates (i, j) of the path result in a single output with the class number. Using the first general approach, it is found that the MSOM algorithm with path matching

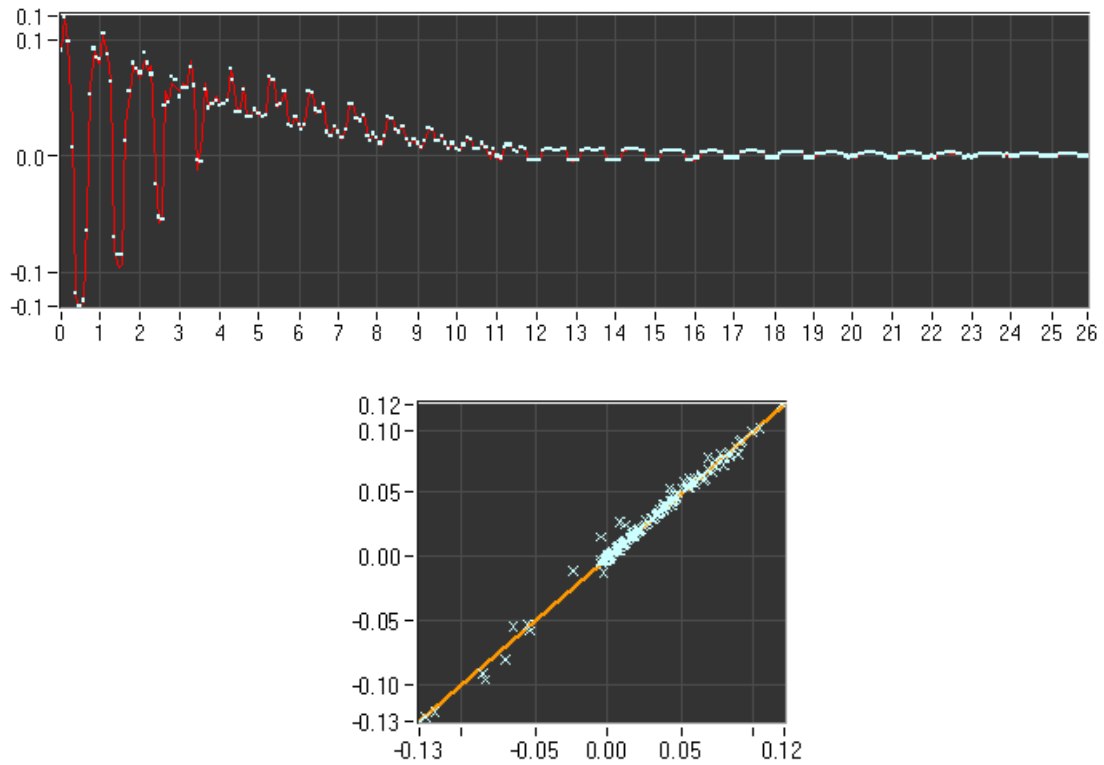


Fig. 51. Activation of the MSOM to a background input pattern (top). Red line is the pattern activated after presenting an input test (blue squares) for a background spatiotemporal response. Plot of the activated pattern values as a function of the input pattern values (bottom).

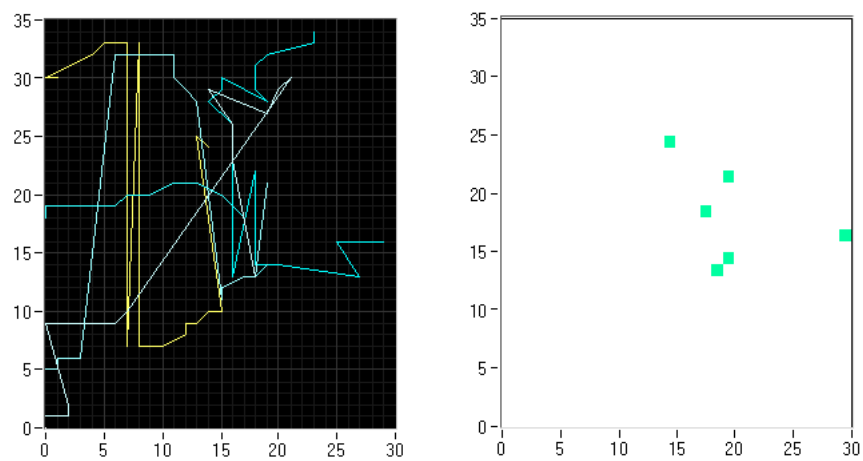


Fig. 52. MSOM activation paths for EM63 spatiotemporal UXO data in orientation C5. Paths (left) and their respective last *bmu* (right).

can provide about 87% correct classification using training inputs and a few test inputs that is about 10% the number of training patterns.

One of the advantages of this method is the possibility to perform classification with a reasonable percentage of success without having to rely on inaccurate background removal using an average of the previous background measurements, or to proceed with an parametric inversion using a simple physics-based model, using background contaminated data. Another advantage about using unsupervised learning with context-based neural maps is the possibility to construct a large database comprising background and target responses in different geological settings and training the MSOM each time new patterns are acquired. In this way, a continuously increasing knowledge base is available to classify target patterns no matter where data are from, with a very fast capacity to provide an answer in real time.

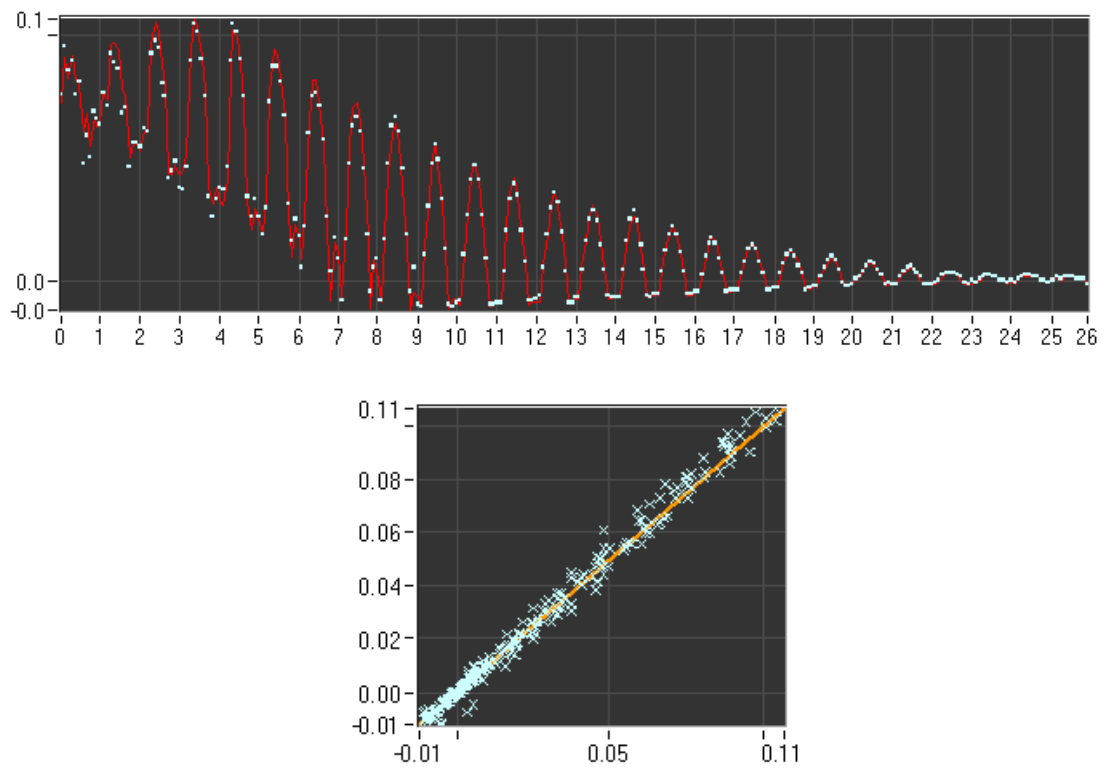


Fig. 53. Activation of the MSOM to a clutter input pattern (top). Red line is the pattern activated after presenting an input test (blue squares) for a background spatiotemporal response. Plot of the activated pattern values as a function of the input pattern values (bottom).

CHAPTER V

CONCLUSIONS

First, this study re-affirms and strengthens the results of Everett and Weiss (2002) that background geological noise in CSEM data exhibits fBm, a fractal signal. This means that spatial fluctuations in the subsurface electrical conductivity structure spanning many length scales contribute to the electromagnetic response in a coherent fashion. Near-surface geological noise can not be averaged out by stacking several profiles along the same line because it is coherent with the transmitter waveform. We have extended the previous published result in this study of one profiles and 2-D maps, with and without presence of metal targets.

The power spectrum of a CSEM profile with a target present does not obey the simple power-law characteristic of a fBm signal. This is because the target signal shifts spectral energy toward higher wavenumbers. The power spectrum can be used only to decide if a target is present, but not to localize it.

Multi-scale analysis techniques such as wavelet transforms are useful methods for target detection and localization. The decomposition of the apparent conductivity profile into approximations and details offer an interesting tool to detect and to locate targets in CSEM profiles. We have shown some promising results for target localization capability in this study.

Using variogram analysis, it is possible to extract spatial conductivity correlations from CSEM data in a fairly easy way. Their use complements FT analysis, and is relevant due to the fractal nature of the near-surface induction response of the geological materials. The profiles and 2-D conductivity maps exhibit spatial correlations that follow a power-law dependence with spatial scale. The variogram trend in the site 3 background conductivity map shows a power-law response at all lag distances

while, in the case of site 2, a test with pipes, the variogram presents a power-law trend for lags smaller than the correlation range and a flat region where the data pairs are no longer correlated. Thus, variogram analysis can also be used to assess the presence of non-geologic targets.

In the case of elongated targets, such as the drill-pipe segments used here, the CSEM response is highly dependent on the Tx-Rx orientation with respect to the target geometry. A promising procedure to enhance target response involves combining responses acquired with orthogonal Tx-Rx orientations at each station. Although this procedure increases by a factor of two the amount of required data, it can remove most of the signals generated by geological noise. The result obtained with a pipe buried close to the ground surface is encouraging, but shows a notable decrease in the S/N ratio due to the increase of distance from the sensors and the attenuation in the conductive host. The use of discrete wavelet filtering in CSEM data has given interesting results for separating the smooth geological components from pipe-like targets. It is important to perform additional tests using more powerful transmitters and the use of multi-frequency, multi-receivers, or time domain instrumentation. Such equipment could be helpful for enhancing target signals and to complement the development of discrimination algorithms based on feature extraction and pattern recognition techniques.

Second, a continuation method for non-linear inversion of spatiotemporal UXO responses has been described. The method reduces the necessity of choosing an initial model that is close to the one that minimizes the cost function. We have shown with four representative examples that continuation path tracking through a low-dimensional model space is numerically stable and can reliably invert spatiotemporal EMI data. Mapping the evolution of model parameters along the path is a useful tool to assess the robustness of the final model and to detect degenerate parameterization

or unimportant parameters. There is a wide range of possibilities to improve the convergence of the current algorithm. These include damping the Hessian matrix inversion or parameterizing the continuation path as a function of arc-length.

The CPU time required for performing the inversion is in the range of seconds, so it may be possible to implement this algorithm in near-real time directly on a field data acquisition platform. The automated nature of the inversion hints at the possibility that, in future, a specifically trained instrument operator may not be required to make the final dig/no-dig decision.

Finally, the application of the paradigm of unsupervised self-organizing maps is adapted for attempting CSEM data clustering and classification. It was demonstrated, by synthetic and field data, that unsupervised learning is able to store a large amount of information about the input datasets, providing data clustering and pathways for classification. The use of SOM for central-loop CSEM transients shows potential capability to perform classification to discriminate background and clutter patterns from, for instance, unexploded ordnance. Finally, with the implementation of a merge SOM algorithm, clustering and classification of spatiotemporal CSEM data was shown to be possible. This is important because the ability to extract target signal from a background-contaminated spatiotemporal pattern is largely desired to avoid the necessity to deal with forward models containing subsurface response or to implement processing algorithm to remove at some degree the effects of background response and the target-host electromagnetic interactions.

REFERENCES

- Ahl, A., 2003, Automatic 1D inversion of multifrequency airborne electromagnetic data with artificial neural networks: discussion and a case study: *Geophysical Prospecting*, **51**, no. 2, 89-98.
- Al-Nuaimy, W., Huang, Y., Nakhkash, M., Fang, M. T. C., Nguyen, V. T., and Eriksen, A., 2000, Automatic detection of buried utilities and solid objects with GPR using neural networks and pattern recognition: *Journal of Applied Geophysics*, **43**, no. 2-4, 157-165.
- An, P., Moon, W. M., and Kalantzis, F., 2001, Reservoir characterization using seismic waveform and feedforward neural networks: *Geophysics*, **66**, no. 5, 1450-1456.
- Ao, C. O., Braunisch, H., O'Neill, K., and Kong, J. A., 2002, Quasi-magnetic solution for a conducting and permeable spheroid with arbitrary excitation: *IEEE Trans. Geosci. Rem. Sens.*, **40**, no. 4, 887-897.
- Aristodemou, E., Pain, C., de Oliveira, C., Goddard, T., and Harris, C., 2005, Inversion of nuclear well-logging data using neural networks: *Geophysical Prospecting*, **53**, no. 1, 103-120.
- Ashida, Y., 1996, Data processing of reflection seismic data by use of neural network: *Journal of Applied Geophysics*, **35**, no. 2-3, 89-98.
- Baçaõ, F., Lobo, V., and Painho, M., 2005, The self-organizing map, the Geo-SOM, and relevant variants for geosciences: *Computers and Geosciences*, **31**, no. 2, 155-163.
- Barreto, G., Araujo, A., and Kremer, S. C., 2003, A taxonomy for spatiotemporal connectionist networks revisited: the unsupervised case: *Neural Computation*, **15**, no. 6, 1255-1320.

Basu, T., Claverie, M., Nolan, D., Yahya, K. B., Suleiman, M., and Carigali, P., 2004, Facies analysis: integration of core and log data using a neural network as input for reservoir modeling in Betty field, Malaysia: *The Leading Edge*, **23**, no. 8, 794-797.

Baum, C. E., 1999, *Detection and identification of visually obscured targets*: Taylor & Francis.

Benaouda, D., Wadge, G., Whitmarsh, R. B., Rothwell, R. G., and MacLeod, C., 1999, Inferring the lithology of borehole rocks by applying neural network classifiers to downhole logs: an example from the ocean drilling program: *Geophysical Journal International*, **136**, no. 2, 477-491.

Benavides, A., and Everett, M. E., 2002, Geological noise and target signals in controlled source electromagnetic data: *SEG 72st Ann. Internat. Mtg*, 1598-1601.

Bescoby, D. J., Cawley, G. C., and Chroston, P. N., 2006, Enhanced interpretation of magnetic survey data from archaeological sites using artificial neural networks: *Geophysics*, **71**, no. 5, H45-H53.

Bhatt, A., and Helle, B., 2002, Committee neural networks for porosity and permeability prediction from well logs: *Geophysical Prospecting*, **50**, no. 6, 645-660.

Boadu, F. K., 1998, Inversion of fracture density from field seismic velocities using artificial neural networks: *Geophysics*, **63**, no. 2, 534-545.

Brown, W. M., Gedeon, T. D., Groves, D. I., and Barnes, R. G., 2000, Artificial neural networks: a new method for mineral prospectivity mapping: *Australian Journal of Earth Sciences*, **47**, no. 4, 757-770.

Bube, K. P., and Langan, R. T., 1994, A continuation approach to regularization for traveltimes tomography: *SEG 64th Ann. Internat. Mtg*, 980-983.

Butler, D. K., 2003, Implications of magnetic backgrounds for unexploded ordnance detection: *Journal of Applied Geophysics*, **54**, 111-125.

Caianello, E., 1961, Outline of a theory of thought-processes and thinking machines: *Journal of Theoretical Biology*, **2**, 204-235.

Calderón-Macías, C., Sen, M. K., and Stoffa, P. L., 2000, Artificial neural networks for parameter estimation in geophysics: *Geophysical Prospecting*, **48**, no. 1, 21-47.

Carr, J. R., 1995, *Numerical analysis for the geological sciences*: Prentice Hall.

Castro de Matos, M., Manassi Osorio, P. L., and Schroeder Johann, P. R., 2007, Unsupervised seismic facies analysis using wavelet transform and self-organizing maps: *Geophysics*, **72**, no. 1, 19-21.

Chamberlin, R. V., Mozurkewich, G., and Orbach, R., 1984, Time delay of the remanent magnetization in spin-glasses: *Phys. Rev. Lett.*, **52**, no. 10, 867-870.

Chappel, G. J., and Taylor, J. G., 1993, The temporal Kohonen map: *Neural Networks*, **6**, 441-445.

Chesney, R. H., Das, Y., McFee, J. E., and Ito, M. R., 1984, Identification of metallic spheroids by classification of their electromagnetic induction responses: *IEEE Trans. Pattern Analysis*, **6**, 809-820.

Christodoulou, C., and Georgiopoulos, M., 2001, *Applications of neural networks in electromagnetics*: Artech House, Boston.

Coppola, E. A., McLane, C. F., Poulton, M., Szidarovszky, F., and Magelky, R. D., 2005a, Predicting conductance due to upconing using neural networks: *Ground Water*, **43**, no. 6, 827-836.

Coppola, E. A., Rana, A. J., Poulton, M., Szidarovszky, F. R., and Uhl, V. W., 2005b, A neural network model for predicting aquifer water level elevations: *Ground Water*, **43**, no. 2, 231-241.

Das, Y., McFee, J. E., Toews, J., and Stuart, G. C., 1990, Analysis of an electromagnetic induction detector for real-time location of buried objects: *IEEE Trans. on Geosci. Remote Sensing*, **28**, no. 3, 278-287.

Daubechies, I., 1992, Ten lectures on wavelets: volume 61 Society of Industrial Applied Math.

Dowd, P. A., and Pardo-Iguzquiza, E., 2005, Estimating the boundary surface between geologic formations from 3D seismic data using neural networks and geostatistics: *Geophysics*, **70**, no. 1, 1-11.

El-Qady, G., and Ushijima, K., 2001, Inversion of DC resistivity data using neural networks: *Geophysical Prospecting*, **49**, no. 4, 417-430.

Erwin, E., Obermeyer, K., and Schulten, K., 1992, Self-organizing maps: ordering, convergence properties and energy functions: *Biological cybernetics*, **67**, 47-55.

Essenreiter, R., Karrenbach, M., and Treitel, S., 2001, Identification and classification of multiple reflections with self-organizing maps: *Geophysical Prospecting*, **49**, no. 3, 341-352.

Everett, M. E., 1996, Homotopy, polynomial equations, gross boundary data, and small Helmholtz systems: *J. Comput. Phys.*, **124**, 431-441.

Everett, M. E., and Weiss, C. J., 2002, Geological noise in near-surface electromagnetic induction data: *Geophysical Research Letters*, **29**, 809-820.

Everett, M. E., and Meju, M. A., 2003, Near-surface controlled-source electromagnetic induction: background and recent advances, *in* Y. Rubin and S. Hubbard, eds., hydrogeophysics: Kluwer, NATO ASI.

Everett, M. E., Benavides, A., and Pierce Jr., C. J., 2005, An experimental study of the time-domain electromagnetic response of a buried conductive plate: *Geophysics*, **70**, G1-G7.

Fedorenko, Y. V., Husebye, E. S., Heincke, B., and Ruud, B. O., 1998, Recognizing explosion sites without seismogram readings: neural network analysis of envelope transformed multistation SP recordings 3-6 Hz: *Geophysical Journal International*, **133**, no. 1, F1-F6.

Fitzgerald, E. M., and Bean, C. J., 2001, Sub-basalt imaging problems and the application of artificial neural networks: *Journal of Applied Geophysics*, **48**, no. 4, 183-197.

Fitzgerald, E. M., Bean, C. J., and Reilly, R., 1999, Fracture-frequency prediction from borehole wireline logs using artificial neural networks: *Geophysical Prospecting*, **47**, no. 6, 1031-1044.

Fuller, B. D., 1971, Electromagnetic response of a conductive sphere surrounded by a conductive shell: *Geophysics*, **36**, no. 1, 9-24.

Garcia, C. B., and Zangwill, W. I., 1981, Pathways to solutions, fixed points and equilibria: Prentice Hall.

Geonics Ltd., EM61 and EM63 metal detectors for unexploded ordnance (UXO) detection and characterization: Geonics, Ltd., Mississauga, Ca, 2001. Technical Note.

Geonics, Ltd., EM63 full time domain electromagnetic UXO detector: Geonics, Ltd., Mississauga, Ca, 2002. Operating Instructions.

Goutorbe, B., Lucazeau, F., and Bonneville, A., 2006, Using neural networks to predict thermal conductivity from geophysical well logs: *Geophysical Journal International*, **166**, no. 1, 115-125.

Guyon, I., 1991, Neural networks and applications tutorial: *Physics Reports*, **207**, no. 3-5, 215-259.

Hammer, B., Micheli, A., Sperduti, A., and Strickert, M., 2004, A general framework for unsupervised processing of structured data: *Neurocomputing* **57**, 57, 335.

Hammer, B., Micheli, A., Neubauer, N., Sperduti, A., and Strickert, M., 2005, Self-organizing maps for time series: *Proceedings of WSOM '05*. September 5-8, Paris, France, 115-122.

Haykin, S., 1994, *Neural networks*: Macmillan Publishing Co., 1st edition.

Helle, H. B., Bhatt, A., and Ursin, B., 2001, Porosity and permeability prediction from wireline logs using artificial neural networks: a North Sea case study: *Geophysical Prospecting*, **49**, no. 4, 431-444.

Herzfeld, C., and Overbeck, C., 1999, Analysis and simulation of scale-dependent fractal surfaces with application to seafloor morphology: *Computers and Geosciences*, **25**, 979-1007.

Hornik, K., Stinchcombe, M., and White, H., 1989, Multilayer feedforward networks are universal approximators: *Neural Networks*, **2**, 359-366.

Hsien-Cheng, C., Kopaska-Merkelb, D. C., and Hui-Chuan, C., 2002, Identification of lithofacies using Kohonen self-organizing maps: *Computers and Geosciences*, **28**, 223-229.

Huang, H. P., and Won, I. J., 2003, Characterization of UXO-like targets using broadband electromagnetic induction sensors: *IEEE Trans. Geosci. Remote Sens.*, **41**, 652-663.

Istratov, A. A., and Vyvenko, O., 1999, Exponential analysis in physical phenomena: *Rev. Sci. Instr.*, **70**, no. 2, 1233-1257.

Jegen, M. D., Everett, M., and Schultz, A., 2001, Using homotopy to invert geo-physical data: *Geophysics*, **66**, no. 6, 1749-1760.

Johnson, V. M., and Rogers, L., 1995, Location analysis in ground-water remediation using neural networks: *Ground Water*, **33**, no. 5, 749-758.

Kangas, J., 1990, Time-delayed self-organized maps: *International Joint Conference on Neural Networks*, New York: IEEE, *International Joint Conference on Neural Networks*, II331-II336.

Katz, A. J., and Thompson, A. H., 1985, Fractal sandstone pores: implications for conductivity and pore formation: *Physical review Letters*, **54**, no. 12, 1325-1328.

Kaufman, A., 1978, Frequency and transient responses of electromagnetic fields created by currents in confined conductors: *Geophysics*, **43**, no. 5, 1002-1010.

Keller, H. B., and Perozzi, D. J., 1983, Fast seismic ray tracing: *SIAM J. Appl. Math.*, **43**, 981-992.

Klose, C. D., 2006, Self-organizing maps for geoscientific data analysis: geological interpretation of multidimensional geophysical data: *Computers and Geosciences*, **10**, 265-277.

Koekkoek, E. J. W., and Booltink, H., 1999, Neural network models to predict soil water retention: *European Journal of Soil Science*, **50**, no. 3, 489-495.

Kohonen, T., 2000, Self-organizing maps: Springer, Berlin, 3rd edition.

Laherrère, J., and Sornette, D., 1998, Stretched exponential distributions in nature and economy: "fat tails" with characteristic scales: *Eur. Phys. J. B*, **2**, 525-539.

Ligtenberg, J. H., 2003, Unravelling the petroleum system by enhancing fluid migration paths in seismic data using a neural network based pattern recognition technique: *Geofluids*, **3**, no. 4, 255-261.

Lin, Z., Fortier, A., and Bartel, D. C., 2002, Classification of salt-contaminated velocities with self-organizing map neural network: *The Leading Edge*, **21**, 1193-1196.

Lo, T., Leung, H., Litva, J., and Haykin, S., 1993, Fractal characterization of sea scattered signals and detection of sea-surface targets: *IEE Proc. F*, **140**, 243-250.

Low, A., Bentin, S., Rockstroh, B., Silberman, Y., Gomolla, A., Cohen, R., and Elbert, T., 2003, Semantic categorization in the human brain: Spatiotemporal dynamics revealed by magneto encephalography: *Psychological Science*, **14**, no. 4, 367-372.

Madsen, K., Nielsen, H. B., and Tingleff, O., Methods for non-linear least squares problems: Technical Report DTU report IMM-REP-2004-01, Technical University of Denmark, 2004.

Mallat, S., 1998, A wavelet tour of signal processing: Academic Press.

Mandelbrot, B., 1983, The fractal geometry of nature: Freeman, San Francisco.

Manoj, C., and Nagarajan, N., 2003, The application of artificial neural networks to magnetotelluric time-series analysis: *Geophysical Journal International*, **153**, no. 2, 409-423.

McCormack, M. D., Zaucha, D. E., and Dushek, D. W., 1993, First-break refraction event picking and seismic data trace editing using neural networks: *Geophysics*, **58**, no. 1, 67-78.

McCulloch, W., and Pitts, W., 1943, A logical calculus of ideas immanent in nervous activity: *Bulletin of Mathematical Biophysics*, **5**, 115-133.

McNeill, J. D., 1980, Electromagnetic terrain conductivity at low induction numbers: Geonics, Ltd., Technical Note TN-6.

____ 1990, Use of electromagnetic methods for underground studies, *in* S. Ward, ed., *Geotechnical and Environmental Geophysics: Society of Exploration Geophysics, Investigations in Geophysics*, 191-218.

McNeill, J. D., and Bosnar, M., 2000, Application of TDEM techniques to metal detection and discrimination: A case history with new Geonics EM-63 fully time-domain detector: Geonics, Ltd., Technical Note TN-32.

Mercado Herrera, V., Russell, B., and Flores, A., 2006, Neural networks in reservoir characterization: *The Leading Edge*, **25**, no. 4, 402-411.

Mikhailov, V., Galdeano, A., Diament, M., Gvishiani, A., Agayan, S., Bogoutdinov, S., Graeva, E., and Sailhac, P., 2003, Application of artificial intelligence for Euler solutions clustering: *Geophysics*, **68**, no. 1, 168-180.

Moutarde, F., and Ultsch, A., 2005, U*F clustering: a new performant cluster mining method based on segmentation of self-organizing maps: *Proceedings of WSOM '05*, September 5-8, Paris, France, 25-32.

Mrozynsky, G., 1998, Analytical determination of eddy currents in a hollow sphere excited by an arbitrary dipole: *IEEE Trans. Mag.*, **34**, 3822-3829.

Mukhopadhyay, A., 1999, Spatial estimation of transmissivity using artificial neural network: *Ground Water*, **37**, no. 3, 458-464.

Nabighian, M. N., 1971, Quasi-static transient response of a conducting permeable two-layer sphere in a dipolar field: *Geophysics*, **36**, no. 1, 25-37.

Nath, S. K., Chakraborty, S., Singh, S. K., and Ganguly, N., 1999, Velocity inversion in cross-hole seismic tomography by counter-propagation neural network, genetic algorithm and evolutionary programming techniques: *Geophysical Journal International*, **138**, no. 1, 108-124.

Nobes, D. C., 1996, Troubled waters: environmental applications of electrical and electromagnetic methods: *Surv. Geophys.*, **17**, 393-454.

Nocedal, J., and Wright, S. J., 1999, Numerical optimization: Springer Series in Operations Research Springer-Verlag.

Pasion, L. R., and Oldenburg, D. W., Locating and characterizing unexploded ordnance using time domain electromagnetic induction:, Technical Report ERDC/GSL TR-01-10, US Army Corps of Engineers, 2001.

Pellerin, L., 2002, Applications of electrical and electromagnetic methods for environmental and geotechnical investigations: *Surv. Geophys.*, **23**, 101-132.

Polzlbauer, G., Dittenbach, M., and Rauber, A., 2005, A visualization technique for self-organizing maps with vector fields to obtain the cluster structure at desired levels of detail: *Proceedings of the IJCNN 2005*, Montreal, Canada, 7.

Poulton, M. M., 2001, Computational neural networks for geophysical data processing: Pergamon, New York.

____ 2002, Neural networks as an intelligence amplification tool: A review of applications: *Geophysics*, **67**, no. 3, 979-993.

Poulton, M. M., Sternberg, B. K., and Glass, C. E., 1992a, Location of subsurface targets in geophysical data using neural networks: *Geophysics*, **57**, no. 12, 1534-1544.

____ 1992b, Neural network pattern recognition of subsurface EM images: *Journal of Applied Geophysics*, **29**, no. 1, 21-36.

Qian, W., and Boerner, D. E., 1995, Electromagnetic modeling of buried line conductors using an integral equation: *Geophys. J. Int.*, **121**, 203-214.

Rheinboldt, W. C., 1980, Solution fields of non-linear equation and continuation methods: *SIAM J. Numer. Anal.*, **17**, 221-237.

Robinson, J. C., 1972, Computer-designed Wiener filters for seismic data: *Geophysics*, **37**, 235-259.

Rosenblatt, F., 1958, The perceptron; a probabilistic model for information storage and organization in the brain: *Psychological Review*, **65**, 386-408.

Russell, B., Ross, C., and Lines, L., 2002, Neural networks and AVO: The Leading Edge, **21**, no. 3, 268-314.

Saggaf, M. M., Toksoz, M. N., and Marhoon, M. I., 2003, Seismic facies classification and identification by competitive neural networks: *Geophysics*, **68**, no. 6, 1984-1999.

Saslow, W. M., 1991, How a superconductor supports a magnet, how magnetically "soft" iron attracts a magnet, and eddy currents for the uninitiated: *Am. J. Phys.*, **59**, no. 1, 16-25.

Schmitz, G. H., Puhlmann, H., Droge, W., and Lennartz, F., 2005, Artificial neural networks for estimating soil hydraulic parameters from dynamic flow experiments: *European Journal of Soil Science*, **56**, no. 1, 19-30.

Seaborne, T. R., Siricki, D. H., and Sowerbutts, W. T. C., 1979, Examples of horizontal loop electromagnetic anomalies controlled by geological faulting: *Geoexploration*, **17**, 77-87.

Singh, S. K., 1973, Electromagnetic transient response of a conducting sphere embedded in a conductive medium: *Geophysics*, **38**, no. 5, 864-893.

Snyder, D. D., MacInnes, S. C., and Hare, J. L., A fast 4-D TEM system for UXO characterization: ESTCP project Final Report 200105, Zonge Engineering & Research Organization, Inc., 2003.

Spichak, V., and Popova, I., 2000, Artificial neural network inversion of magnetotelluric data in terms of three-dimensional earth macroparameters: *Geophysical Journal International*, **142**, no. 1, 15-26.

Spichak, V., Fukuoka, K., Kobayashi, T., Mogi, T., Popova, I., and Shima, H., 2002, ANN reconstruction of geoelectrical parameters of the Minou fault zone by scalar CSAMT data: *Journal of Applied Geophysics*, **49**, no. 1-2, 75-90.

Stalnaker, J. L., Everett, M. E., Benavides, A., and Pierce Jr., C. J., 2006, Mutual induction and the effect of host conductivity on the EM induction response of buried plate targets using 3-D finite-element analysis: *IEEE Transactions on Geoscience and Remote Sensing*, **44**, no. 2, 251-259.

Strecker, U., and Uden, R., 2002, Data mining of 3D poststack seismic attribute volumes using Kohonen self-organizing maps: *The Leading Edge* **21**, 1032-1037.

Strickert, M., and Hammer, B., 2005, Merge SOM for temporal data: *Neurocomputing*, **64**, 39-71.

Tezkan, B., 1999, A review of environmental quasi-stationary electromagnetic techniques: *Surv. Geophys.*, **20**, 279-308.

Tingdahl, K. M., and de Rooij, M., 2005, Semi-automatic detection of faults in 3D seismic data: *Geophysical Prospecting*, **53**, no. 4, 533-542.

Torrence, C., and Compo, G. P., 1998, A practical guide to wavelet analysis: *Bull. Am. Meteor. Soc.*, **79**, 61-78.

Trappe, H., and Hellmich, C., 2000, Using neural networks to predict porosity thickness from 3D seismic: *First Break*, **18**, no. 9, 377-384.

Turcotte, D. L., 1992, *Fractals and chaos in geology and geophysics*: Cambridge University Press.

Turcotte, D. L., and Huang, J., 1995, Fractal distributions in geology, scale invariance, and deterministic chaos, *in* C. C. Barton and P. R. La Pointe, eds., *Fractals in the earth sciences*: Plenum Press, *Investigations in Geophysics*, 1-38.

Ultsch, A., 2003a, Maps for the visualization of high-dimensional data spaces: *Proceedings of WSOM '03*, Fukuoka, Japan, 225-236.

____ 2003b, U*-matrix: a tool to visualize clusters in high dimensional data: Department of Computer Science, University of Marburg, Technical Report 36.

van der Baan, M., and Jutten, C., 2000, Neural networks in geophysical applications: *Geophysics*, **65**, no. 4, 1032-1047.

Varsta, M., Heikkonen, J., and del Ruiz Millán, J., 1997, Context learning with the self organizing maps: Proceedings of WSOM '97, Helsinki, Norway.

Vasco, D. W., 1994, Singularity and branching: A path following formalism for geophysical inverse problems: *Geophys. J. Internat.*, **19**, 809-830.

____ 1998, Regularization and trade-off associated with non-linear geophysical inverse problems: penalty homotopies: *Inverse Problems*, **14**, 1033-1052.

Villiers, N. D., and Glasser, D., 1981, A continuation method for non-linear regression: *SIAM J. Numer. Anal.*, **18**, 1139-1154.

Voegtlin, T., 2002, Recursive self-organizing maps: *Neural Networks*, **15**, no. 8-9, 979-992.

Wang, L., and Mendel, J. M., 1992, Adaptive minimum prediction-error deconvolution and source wavelet estimation using Hopfield neural networks: *Geophysics*, **57**, no. 5, 670-679.

Watson, L. T., 1989, Globally convergent homotopy methods: a tutorial: *Appl. Math. Comp.*, **31**, 369-396.

Weichman, P. B., 2003, Universal early-time response in high-contrast electromagnetic scattering: *Phys. Rev. Lett.*, **91**, no. 14, 143-908.

Wiemer, J. C., 2003, The time-organized map (TOM) algorithm: extending the self-organizing map (SOM) to spatiotemporal signals: *Neural Computation*, **15**, 1143-1171.

Zhang, Y., and Paulson, K. V., 1997, Magnetotelluric inversion using regularized Hopfield neural networks: *Geophysical Prospecting*, **45**, no. 5, 725-743.

Zhang, Z., and Zhou, Z., 2002, Real time quasi-2-D inversion of array resistivity logging data using neural network: *Geophysics*, **67**, no. 2, 517-524.

Zhang, L., Fortier, A., and Bartel, D. C., 2002a, Interpreter's corner-classification of salt-contaminated velocities with self-organizing map neural network: *The Leading Edge*, **21**, no. 12, 1193-1196.

Zhang, L., Poulton, M. M., and Wang, T. 2002b, Borehole electrical resistivity modeling using neural networks: *Geophysics*, **67**, no. 6, 1790-1797.

Zhang, Y., Collins, L., and Carin, L., 2003, Physics model based unexploded ordnance discrimination using wideband EMI data: *Detection and remediation technologies for mines and minelike targets VIII*, 1023-1045.

APPENDIX A

GEONICS EM63 TECHNICAL DATA

Geonics EM63 is a state-of-the-art metal detector; a compact version of modern time-domain geophysical instruments used for oil and mining exploration. Originally suited for unexploded ordnance detection, the EM63 can also be used for near-surface conductivity mapping and all-purpose metal detection. There are two main components: transmitter and receiver. Both are controlled by a solid-state field computer (Juniper systems). The computer uses a PCMCIA multi-purpose DAQ card for control and data acquisition. All the system is powered by 12 *v* Pb-Gel acid battery.

Transmitter Data

- 1 \times 1 *m* multi-turn coil
- 16 *A* maximum current intensity
- 500 *A/m* maximum magnetic field intensity
- bipolar fin-shaped square waveform
- housing: fiberglass

Receiver Data

- Three 0.5 \times 0.5 *m* multi-turn coils
- Bottom coil: main receiver; co-planar with TX coil
- Middle coil: secondary coil (one time channel acquisition)
- Top coil: compensation coil (common noise rejection)

Controller Data

- Computer Juniper Systems PRO 4000
- 80486DX CPU using a DOS operating system, PCMCIA DAQ card
- Analog: 8 inputs, 2 outputs
- Resolution 16 bits
- Sampling rate 100 ks/s
- Digital: 24 I/O lines
- experimental noise floor: $20\text{ }\mu\text{v}$

Fin-shaped square waveform is a on:off:-on:off sequence. The *on* waveform consist of a “slow” rise with characteristic time of 4.3 ms . After reaching the maximum stationary current intensity, current if ramp-off (time not specified but shorter than 0.1 ms). This fast ramp-off of the current is used to provide the electromotive force to induce currents in the subsurface and targets. These induced currents gives rise to a secondary field that is measured by the receiver coils. The transmitter driver can generate two waveform patterns with different periods (called *hi* and *medium*) to accommodate two different data acquisition rates.

Receiver coil voltages induced on the bottom and mid coils produced are conditioned by filters and a pre-amplifier stage before being digitized. These voltages are stored in binary format using a set of 26 log-spaced time channels (in the medium speed transmitter mode) as described on table 1. Data can be stacked to improve the signal-to-noise ratio. Manufacturer claims that stacked data can provide 18-bits of resolution.

Table 1. Geonics EM63 time channels data.

Timechannel	Time (<i>ms</i>)	Timechannel	Time (<i>ms</i>)
1	0.177	14	1.714
2	0.191	15	2.157
3	0.216	16	2.724
4	0.246	17	3.445
5	0.286	18	4.361
6	0.336	19	5.535
7	0.400	20	7.032
8	0.480	21	8.858
9	0.584	22	10.720
10	0.714	23	13.150
11	0.883	24	16.200
12	1.097	25	20.050
13	1.366	26	25.010

VITA

GENERAL INFORMATION

Name: Alfonso Benavides Iglesias

Home Address: La Salle Ave. Pancho Building 3-22. Los Caobos, Caracas 1050
Venezuela

Home Phone: +1-212-793-9134.

e-mail: a0b9022@neo.tamu.edu, benalf@cantv.net

ACADEMIC INFORMATION

Ph.D. in Geophysics, Texas A & M University, May 2007

M.A.S., Universidad Central de Venezuela, Caracas, September 1999

B.S., Universidad Central de Venezuela, Caracas, September 1994

(Graduation: Dec, 1994)